LONG ISLAND SOUND STUDY
EPA ASSISTANCE AWARD
FINAL REPORT

May 25th, 2010

1. **EPA Grant Number:** LI-97127101 **Project Title:** Simulation of Long Island Sound with the System-Wide Eutrophication Model (SWEM): Inter-annual Variability and Sensitivity

2. **Grantee Organization:** University of Connecticut

3. **Contact Names:** Dr. James O'Donnell, Dr. Hans Dam, Grant McCardell and Todd Fake.

4. Public Summary

Very low levels of dissolved oxygen (DO) have occurred in the near bottom waters of western Long Island Sound (LIS) each summer since measurements began in 1988 and this causes stress to the marine life in the region. Similar patterns of low DO occur in many urbanized estuaries and the phenomenon is termed hypoxia. To mitigate the extent and duration of hypoxia in LIS, the EPA and the States of New York and Connecticut have developed a Comprehensive Conservation and Management Plan (CCMP) using a computer model to assess the likely impact of reductions in nitrogen discharged from water treatment plants and non-point sources. An improved model, the System Wide Eutrophication Model (SWEM), has been developed and tested by Hydroqual Inc. to simulate the biogeochemistry and circulation in Long Island Sound and adjacent waters and is being used to reassess the effectiveness of the plan. SWEM is a complex model with many parameters that represent the rates of the processes that influence the DO concentration. In this project we implemented SWEM at the University of Connecticut to independently study the sensitivity of the model predictions of DO concentrations to parameter choices and boundary conditions, and to assess the effect of year-to-year variations in precipitation and wind patterns. Since there have been many important advances in the understanding of LIS and the coastal ocean since the initiation of the development of SWEM, we also assess whether SWEM is consistent with recent new knowledge.

SWEM has two major modules. Circulation and mixing are simulated by solving the equations describing the hydrodynamics of the coastal ocean with boundary conditions that represent river flow, winds and the state of the ocean at the model boundaries. This component is called ECOM. The products of this module (velocities and vertical eddy coefficients) are passed to the water quality module to compute the evolution of nutrients, plankton, dissolved oxygen etc. During the development and calibration of SWEM, an ad-hoc reduction to the vertical eddy coefficients predicted by the ECOM module was introduced to reduce near-bottom dissolved oxygen in the western Sound in the summer. Recent work on mixing in the coastal ocean and comparison of ECOM results to recent observations in the Sound suggest that the original ECOM values were actually realistic and that the values imposed by the ad-hoc reductions were much too small. The need to distort the representation of mixing in order to obtain a reasonable representation of hypoxia points to other under-lying problems in SWEM. By comparing recent observations in LIS and SWEM predictions, we found that both respiration and production were significantly underestimated in SWEM while the levels of mixing predicted by the original circulation module were in a realistic range. We conclude that it is likely that the underestimation of respiration was the cause of the problems that led to the need for the artificial reduction of vertical mixing rates. Though outside the scope of the original project, we investigated the cause of the under-estimation of respiration and found we could correct it and improve the model skill without the distortion of vertical mixing that is currently implemented in the SWEM model.

To quantitatively assess how well the model performed we employed a statistical summary of the misfit between the model solutions for 1988-99 and 1994-95 with observations of the DO concentrations made by ship surveys in these same years. We called how well the model matched the actual observed data the model *skill*. We used several variants of this metric, but for the purposes of comparing different versions of the same model, a higher skill means a better simulation. SWEM has 120 parameter values that must be selected. We examined the effect on the model skill of increasing and decreasing each parameter by 10% from the value chosen by Hydroqual for simulations of 1988-99 and 1994-95. The results demonstrated which parameters were most important and suggested some changes to parameter values that might lead to improved predictions.

We also investigated the consequences of doubling the fluxes of nitrogen from point sources and setting these to zero on the predictions of the model. We found little difference between the solutions in areas prone to hypoxia. With zero discharges, the minimum DO in the late summer was approximately 1mg/L higher than with using the 1995 discharge levels and only 0.1 mg/L lower with double the 1995 discharges. Since SWEM was designed to mimic the existing plankton community structure, which would be expected to be different if nutrient ratios were changed radically, these simulations should not be considered realistic. However, they do provide a bound on the magnitude of DO changes to be expected from management actions which might reduce discharges 30 to 50%. The model predicts that changes are likely to be small. Of course this result must be checked once production and respiration rates are simulated more realistically.

The validation and verification studies of SWEM used two years of observations. However, it is well established that there are significant differences in the pattern of river discharge and wind variations from year to year, so it was important to assess whether these simulation years were representative. We contracted with Hydroqual to provide ECOM solutions for four additional years (1998-99, 1999-2000, 2000-2001, and 2001-2002) and used these, together with the 1994-95 boundary source fluxes, to simulate what the conditions in LIS would have been with different weather conditions. We then compared these solutions and found that they were similar to each other. We conclude that SWEM does not have the ability to distinguish the influence of inter-annual meteorological variability on hypoxia.

We finish our study with recommendation for future research needs. Clearly, the model must be improved to better represent the critical processes that determine the rate of decline in DO in the summer: production, respiration, and vertical mixing. We have outlined how that should be accomplished. This is a substantial change and the implications for the management decisions that have been instituted should be reassessed by repeating at least some of the simulations used in the development of total maximum daily loads (TMDLs). However, that is not enough. For the model to be useful in predicting the impact of changes to sources of nutrients, it is essential to observe how community respiration and production rates depend upon the available nutrient concentrations in LIS. A capability to simulate that response will establish confidence that the model can represent the effect of management actions correctly. This will require sustained field observations in addition to the existing monitoring program.

Recent buoy observations have shown that the magnitude of the variation in DO and chlorophyll that occurs during a single day is comparable to the differences in the monthly mean concentrations between years. It seems likely that nutrients will show similar variability. Since the surveys occur during the day and only twice a month, it is not surprising that comparing this data to the model predictions does not really assess whether the model is adequate. Much of the difference between the predictions and observations is due to the uncertainty in the observation of the several-day mean. Though the precision of the measurements may be high, the sampling scheme does not provide a good estimate of the values that SWEM was being used to predict.

Further, only the availability of high frequency moored DO observations and ship based measurements of respiration and production rates provided the insight that led to identification of the weaknesses in the model. The existing observations of the community respiration and production rates show that these are highly variable in time and location. It is unclear what causes this variability. The high frequency variability in the nutrient levels are unknown at present and this must be established. Finally, the crude resolution of the lateral variations in the circulation model essentially prohibit the possibility that lateral gradients strong enough to drive substantial lateral circulation can be established. The low spatial resolution of the model also diminishes its value to the management of water quality in embayments like Hempstead Harbor, Smithtown Bay, etc.

We believe that some improvements to the modeling process are also needed. It is critical that engagement between observationalists and theoreticians be enhanced. There are good paradigms for how to do this. Use of open-source models and data sharing standards are essential to allow more engagement by experts and constructive criticism.

5. **Grant Period:** 1 October 2005 to 30 May 2009

6. **Project Description:**

The System-Wide Eutrophication Model (SWEM) has been developed to simulate the circulation and dissolved oxygen dynamics of Long Island Sound (LIS) and adjacent waters. It has been applied to assist in the design and evaluation of the LIS Comprehensive Management Plan (CCMP). However, the evaluation of the model's veracity and its utility as a tool for managers has been limited by the availability of data and computing power. The lack of independent confirmation of the model predictions has raised concerns about the accuracy of such predictions. Altogether, these factors have hindered the development of a broad consensus for action.

The primary objectives of this project are:
(1) To establish an independent, quantitative evaluation of the effectiveness of the SWEM model in the simulation of the water quality and circulation of Long Island Sound, and
(2) To identify and prioritize additional studies that will improve our ability to predict the impact of management strategies on the water quality of Long Island Sound.

In the conduct of this project we will execute the following tasks:

> **Task 1**. We will obtain **SWEM from Hydroqual and implement the model at the SUN Microsystems Center of Excellence in Oceanographic Computing (CEOC)** in the Department of Marine Sciences and replicate the Hydroqual results for 1989 and 1995.
>
> **Task 2**. We will **establish quantitative measures of the model skill** and summarize the results for 1989 and 1995. To identify the model components that determine the magnitude of these uncertainties, the mean square difference between predicted and observed variables (e.g. salinity, temperature, dissolved oxygen, nitrogen, phosphorous current velocity, chlorophyll) are more useful. We will explore the links between our ability to simulate the various subcomponents and the over-all model performance.
>
> **Task 3**. We will **evaluate the transport predictions** of SWEM in 1995 by comparing them to observations.
>
> **Task 4**. To prepare to simulate additional years, we will **collate data** for weather, river discharge, in-situ current and water quality and effluent discharge for 1999-2002 and prepare an on-line data access portal.
>
> **Task 5**. We will **establish the sensitivity of the model** (as indicated by analysis of skill measures) to the choice of model parameters; we will replicate the 89 and 95 simulations in parallel by exploiting the 54 processors available to us in the SUN CEOC. We will then analyze how the measures of model skill depend upon the uncertainty in the various parameters and identify the parameters that need to be better constrained by additional field or lab work.
>
> **Task 6**. In addition to the sensitivity to the choice of model parameters, predictions of water quality are also sensitive to the details of the model formulation. We will, therefore, **explore the sensitivity to circulation model predictions**.

Task 7. We will **perform simulations for the years 1999-2002** using the forcing fields assembled under Task 4 and compute the measures of model skill. We will then perform a parallel sensitivity analysis similar to that undertaken in Task 5.

Task 8. To further **examine the relative importance of inter-annual weather and river discharge variations** to the dissolved oxygen distribution, we propose to conduct numerical experiments in which we hold nutrient loadings at 1995 levels. We will then compute solutions using the weather and stream-flow from 1995 (low discharge) and 2000 (normal discharge). The differences in the solutions will clarify the sensitivity of the ecosystem to external forcing so that is can be compared to the uncertainties due to parameter choices.

Task 9. To fully exploit the investment in the SWEM model we propose to develop and evaluate its ability to **produce short-term (monthly) forecasts** of the extent of the hypoxia in the summer. If these forecasts can be demonstrated to have skill, then they may have value to users of the natural resources for planning and management purposes.

Task 10. In the final phase of the program we will design and perform model experiments that will demonstrate which model parameters/formulations are most important to the determining the performance of the model. Our intent is to **produce a tool that will guide managers of the research program as to how to most effectively prioritize and invest** in studies that will improve our understanding and ability to predict the behavior of the ecosystem.

Task 11. Our project will conclude with a set of broad recommendations for follow-on work that will be needed to address the weakness in our understanding of the LIS ecosystem. These will focus on how to improve the predictive skill of the model.

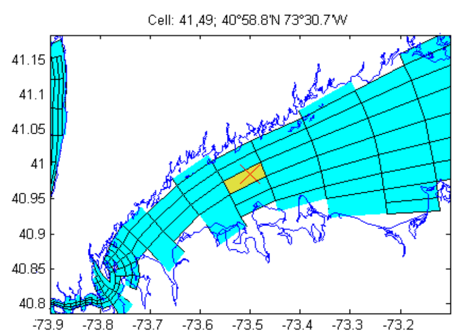
7. **Activities and Accomplishments:**

7.1 Task 1. Installation of Simulation Capability

With the support of the EPA and funds from the University of Connecticut, we have constructed a 48-processor parallel computing system with 2 Terabytes of disk and installed the SWEM code. This capability enables rapid calculations and the simulation of a wide range of model scenarios with which to assess the sensitivity of the predictions to parameter choices. Though the model code does not exploit the parallel processing capability of the infrastructure, we can run 48 alternative scenarios simultaneously and complete a year-long simulation in 4-6 hours. We also built a MATLAB-based system to access the solution files, perform analyses and generate graphics. These routines could be made available or converted to provide access to users remotely via the internet.

To ensure that our implementation of SWEM was performing properly, we replicated the simulations of Hydroqual Inc. for 1995 and demonstrated good agreement. The SWEM grid in western LIS and the location of CTDEP survey program station C2 are indicated in Figure 1(a). Figure (1b) shows a comparison of the evolution of the DO near the surface (blue) and near bottom (green) DO in SWEM for the 1988-1989 simulation at the model cell containing CTDEP station C2. The red symbols show the range of observations of DO observed at the station during the survey cruises. Figure 2 shows the along sound variation of near surface and bottom DO obtained from our 1995 simulation and Figure 3 shows a similar plot from Hydroqual Inc. (2003). These are equivalent.

(a)



(b)

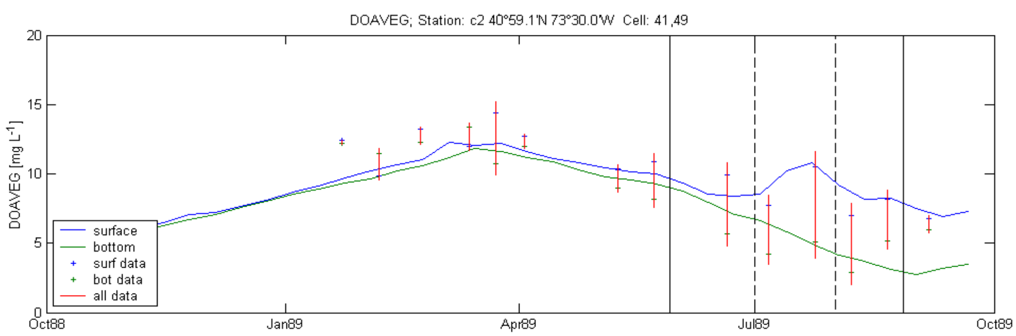


Figure 1(a) map of the coastline of western LIS with the SWEM grid and the location of the CT DEP survey station C2. (b) shows the evolution of the weekly averaged near surface DO (blue) and near bottom DO (green) predicted by SWEM for 1988-1989. The observed concentrations at the surface and bottom are represented by the * and + symbols respectively, and the range of the profile data is shown by the red lines.

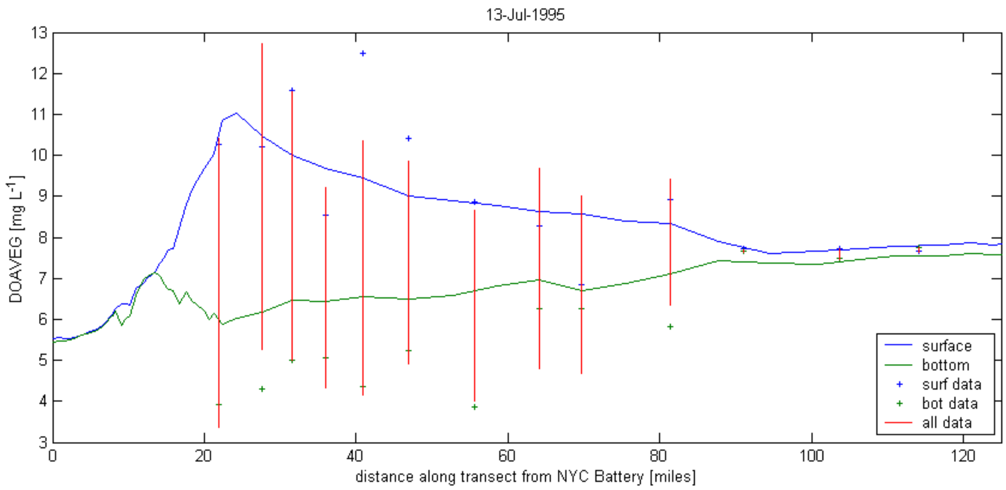


Figure 2. Comparison of the along Sound structure of the near surface (blue) and near bottom (green) DO concentration predicted by SWEM for July 1995. The surface and bottom observation of CTDEP are shown by the * and + symbols respectively, and the range of the profile data is shown by the red lines.

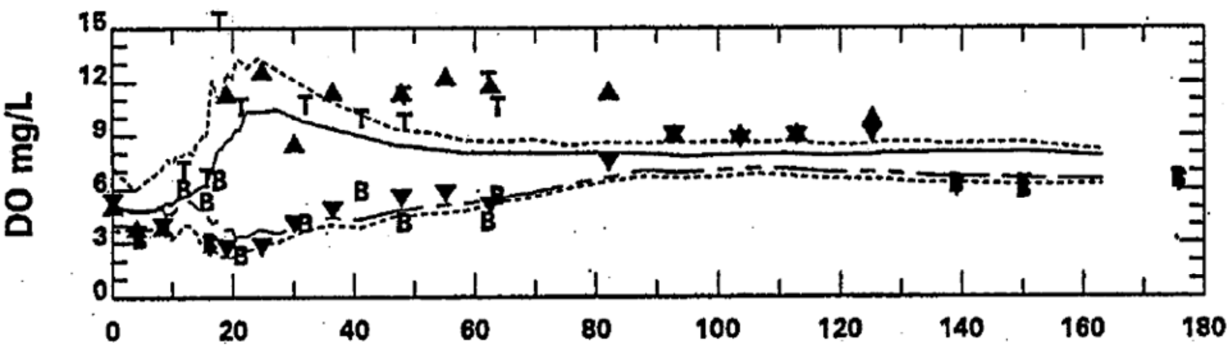


Figure 3. Comparison of the along Sound structure of the near surface (solid line) and near bottom (long dash-short dash line) DO concentration predicted by SWEM for July 1995. The uppermost and lowest dashed lines show the highest and lowest values obtained in the averaging period. The triangles show the survey data along the axis of the Sound and the T symbols show the near shore station observations. From Hydroqual Inc. (2003).

7.2 Task 2. Measures of Model Performance

We examined a variety of approaches to quantifying the difference between model predictions and observations. Hydroqual Inc. (2003) relied mainly on graphical correlation; however, since the model has 25 variables that it predicts at each grid point and we conducted hundreds of simulations in the study of the sensitivity of misfit to parameter choice, a different approach was necessary. The SWEM variables are listed in Table 1.

Table 1. Variables used in SWEM. * Indicates variables not in original output list.

SWEM SYSTEM NUMBER	Biogeochemical Variable	Symbol
1	SALINITY	SAL
2	PHYTOPLANKTON - WINTER DIATOMS	PHYT1
3	PHYTOPLANKTON - SUMMER ASSEMBLAGE	PHYT2
4	PARTICULATE ORGANIC PHOSPHOROUS - REFRACTORY	RPOP
5	PARTICULATE ORGANIC PHOSPHOROUS - LABILE	LPOP
6	DISSOLVED ORGANIC PHOSPHOROUS - REFRACTORY	RDOP
7	DISSOLVED ORGANIC PHOSPHOROUS - LABILE	LDOP
8	TOTAL DISSOLVED INORGANIC PHOSPHOROUS	PO4T
9	PARTICULATE ORGANIC NITROGEN - REFRACTORY	RPON
10	PARTICULATE ORGANIC NITROGEN - LABILE	LPON
11	DISSOLVED ORGANIC NITROGEN - REFRACTORY	RDON
12	DISSOLVED ORGANIC NITROGEN - LABILE	LDON
13	TOTAL AMMONIA	NH3T
14	NITRITE + NITRATE	NO23
15	BIOGENIC SILICA - UNAVAILABLE	BSI
16	TOTAL SILICA - AVAILABLE	SIT
17	PARTICULATE ORGANIC CARBON - REFRACTORY	RPOC
18	PARTICULATE ORGANIC CARBON - LABILE	LPOC
19	DISSOLVED ORGANIC CARBON - REFRACTORY	RDOC
20	DISSOLVED ORGANIC CARBON - LABILE	LDOC
21	DISSOLVED ORGANIC CARBON - REACTIVE	REDOC
22	DISSOLVED ORGANIC CARBON - ALGAL EXUDATE	EXDOC
23	O2* - AQUEOUS SOD	02EQ
24	DISSOLVED OXYGEN	DO
25*	TOTAL ACTIVE METAL	TAM
26*	TEMPERATURE	RCATEMP
27	Bottle BOD including suspended SOD	BOTBODHS
28	Chlorophyll (sum of Phyt1 & 2 time ccratios)	CHLAVEG
29	ECOM salinity	HYDSAL
30*	Gross Primary Production above 1% light level	TGPDP
31*	Ecom Temperature	HYDTEMP
32*	Total Respiration in oxygen units including algal, chemical and sedin	TRESP

We adopted the Brier Skill, see Storch and Zwiers (1999), as our metric of model performance. This is written as

$$S_1 = 1 - \frac{1}{N_p}\sum_{i=1}^{N_p}\frac{1}{N_i}\sum_{j=1}^{N_i}\left(d_{i,j} - p_{i,j}\right)^2 / \frac{1}{N_p}\sum_{i=1}^{N_p}\frac{1}{N_i}\sum_{j=1}^{N_i}\left(d_{i,j} - r_{i,j}\right)^2 \qquad (1)$$

where the vector $d_{i,j}$ contains the $j = 1,2,..N_i$ measurements of the $i = 1,2,...N_p$ variables at the LIS stations. Similarly, $p_{i,j}$ are the model predictions at the same time and location of the data, $d_{i,j}$, and the vector r_{ri} contains the predictions of the reference model. S_1 compares the ratio of the variance in the observations not explained by the model to that not explained by a reference model. Note that Equation (1) expresses S_1 for all variables but it can also be expressed for a single variable, e.g. the dissolved oxygen concentration (DO). If $S_1 > 0$ then the model is in better agreement with the data than the reference model. Conversely, if $S_1 < 0$, it is not as good. Note that complexities in interpretation arise when there are more observations in some months and at some stations than others. However, we have chosen to include only the stations in LIS that are sampled most frequently. The lateral structure is not very well resolved in the model and the number of variables in $S_1$is also subject to choice. We will focus on either all variables, or just DO ($N_P = 1$).

We considered several alternative reference models but report only two. The first, the monthly mean observations computed at each station in the CT DEP surveys (see Kaputa and Olson, 1999) between 1987 and 2005, is a very high standard to surpass. The term in the denominator in Equation (1) is the variance associated with the year to year, or inter-annual, variability in properties. We chose this reference model because a model that is to be used for management purposes should be able to simulate both the variability that is observed within one year and distinguish between years. For this reference model a positive skill ($S_1 > 0$) suggests that the model can predict inter-annual differences in the property distributions. We call this model the climatology.

A lesser challenge is provided by the choice of the annual mean at each station as the reference model. The terms in the denominator in Equation (1) are then the variance observed at a station. In this case, $S_1 > 0$ suggests that the model can predict aspects of the annual cycle.

It is important to note that there are always errors in the observations due not only to the precision of the instruments and analyses methods, but also due to the difference in the property simulated (the average over a volume of order $10^7 m^3$ and many days) and that measured (a few samples from a bottle with a volume of order $10^{-1} m^3$). Even a perfect model, therefore, should not be expected to have $S_1 = 1$. Finally, we note that a change to a model (whether in formulation or parameter choices) will lead to a change in S_1 that is positive if the model is a better representation of the observations irrespective of the choice of reference model.

The locations of the data stations used for the skill calculations are shown in Figure 4. At this subset of stations, measurements are available for both evaluation years (1989 and 1995). Table 2 shows the 12 SWEM parameters for which skill calculations were made. These parameters are those which are both represented in SWEM and measured by the LISS in 1988-89 and the CTDEP in 1994-95.

Table 2. SWEM variable and Long Island Sound Study-CTDEP survey data used in the calculation of Overall Skill.

Quantity	Variable Code		
	SWEM	LISS	CTDEP
particulate organic phosphorus	TPOP	TPP	PP-LC
dissolved organic phosphorus	TDOP	DOP	TDP
phosphate	DPO4	PO4	DIP
dissolved organic nitrogen	TDON	DON	TDN-LC
particulate organic carbon	TPOC	POC	PC
dissolved organic carbon	TDOC	DOC	DOC
particulate silica	TPSI	Psi	BIOSI-LC
dissolved silica	DSI	Dsi	SIO2-LC
dissolved oxygen	DO	DO Titra	DO Titra
nitrite plus nitrate	NO23	NO23	NOX-LC
ammonia	DHN3	NH4	NH#-LC
chlorophyll	CHLA	Chla	CHLA

Figure 5a shows the full SWEM computational domain as a square matrix with the ocean shaded in blue and the land shaded green. Note that this evaluation focuses on the model performance within Long Island Sound. The model cells in LIS are enclosed in the red line and form a very small subset of the full SWEM domain. We modified the SWEM code to save all variable values at the cells shaded red to facilitate analysis of the model processes. The geographic structure of the model is shown in Figure 5b.

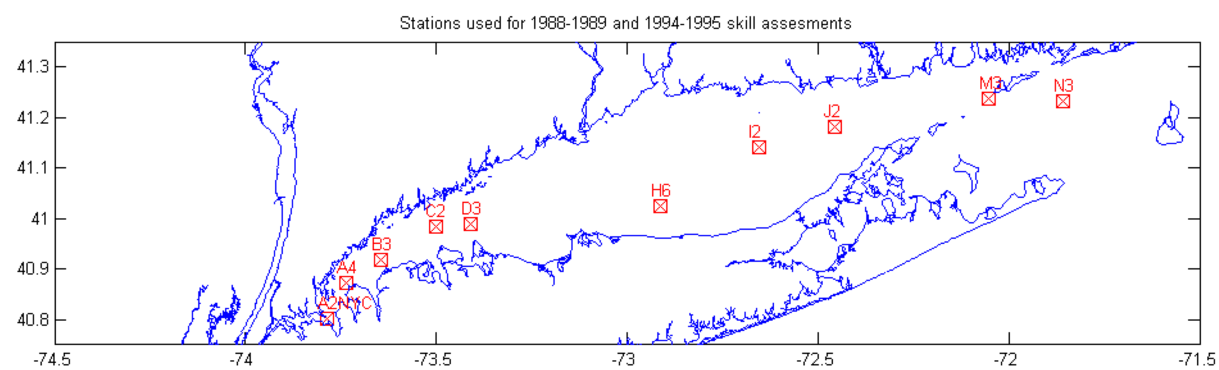


Figure 4. Coastline of Long Island Sound together with the locations of the CT DEP survey locations used in the model evaluation.

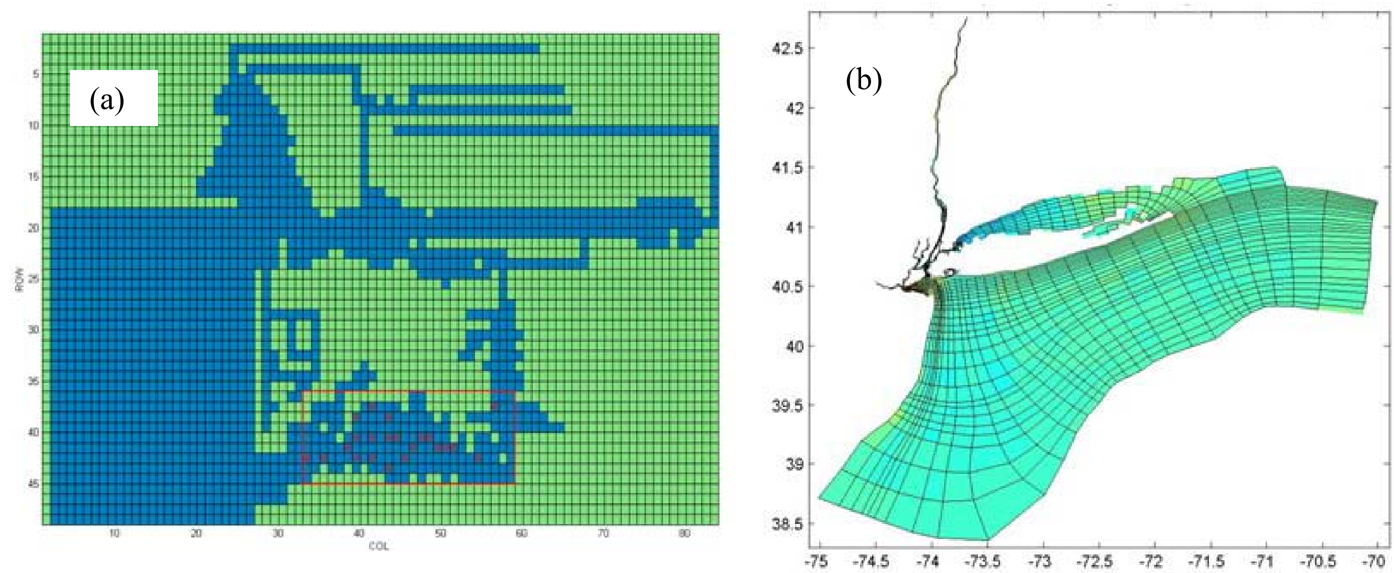


Figure 5. (a) The SWEM domain with the land shaded in green and the ocean shaded blue. The cells that simulate LIS are within the red box. (b) A geographic representation of the geometry of the model grid.

Using the monthly climatology as the reference model we computed $S_1 = -.25$ for 1989 and $S_1 = -2.25$ for 1995. Note that although S_1 is greater in 1989 than in 1995, it is not correct to conclude that the model performed better in 1989 since the climatological reference model will perform better in a typical year than in an atypical one. These are different simulations. However, since $S_1 < 0$ in both cases, predictions made using the climatology were better representations of the observations in both years.

Figure 6 shows a histogram contrasting the skill for selected variables (listed in the Figure caption) in the SWEM simulations of 1989 and 1995. Clearly, there are a wide range of values, reflecting the fact that some variables are better simulated by SWEM than others. It is also evident the variables have different skills in the two years. However, we reiterate that this is partly a consequence of the inter-annual variability in the observations and that it is changes from these values of S_1 that we are concerned about. Note that most are negative indicating that SWEM does not represent the difference between the simulated year and the climatology.

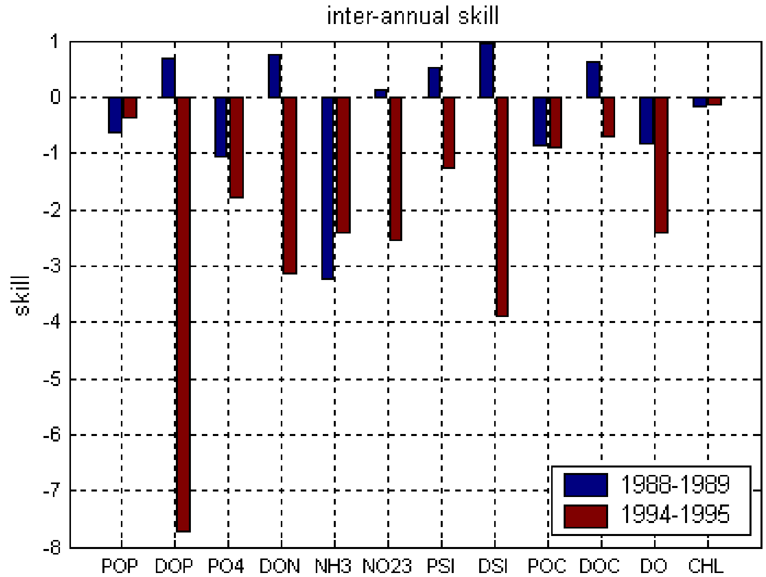


Figure 6. Skill of SWEM predictions for 1989 (blue) and 1995 (red) for Particulate and Dissolved Organic Phosphorus (POP and DOP), Phosphate (PO4), Dissolved Organic Nitrogen (DON), Ammonium (NH3), Nitrate plus Nitrite (NO23), Particulate Silica (PSI), Particulate Organic Carbon (POC), Dissolved Organic Carbon (DON), Dissolved Oxygen, (DO), and Chlorophyll (CHL).

7.3 Task 3. Transport Model Evaluation

The transport module of SWEM is used to predict the patterns of horizontal velocity and its evolution. In addition, it simulates the heat and salt transport. These influence the density and stratification of the estuary and the stratification and vertical gradient in the horizontal velocities then determine the vertical flux of momentum due to turbulence. During the model development, Hydroqual (2003) compared the predictions to observations of sea level, currents, temperature and salinity and showed good agreement. However during the testing of the water quality components it was found necessary to reduce the magnitude of the vertical flux of DO to develop a realistic simulation of the observed degree, duration and extent of hypoxia. Since there have been considerable advances in the understanding of mixing in the coastal ocean in the last decade, a reassessment of the representation of hydrodynamic processes in SWEM is warranted.

The LISICOS program at the University of Connecticut maintained an extensive array of current meters in LIS between 2004 and 2007. This data set, together with data from a few other long term current meter moorings has recently been analyzed by Bennett et al. (2010) to reveal the spatial structure of the tides and influence of seasonal stratification. Figure 7 shows the locations of the data sources. There are 6 stations indicated in the area of the western Sound where hypoxia appears every year. A 7th is not shown because of the scale of the map. The array provided very good resolution of the spatial structure and Figure 8 shows the amplitude and orientation of the near surface and near bottom M2 tidal current ellipses in the western-most stations. Little lateral structure is apparent and this supports the use of coarse resolution in the model.

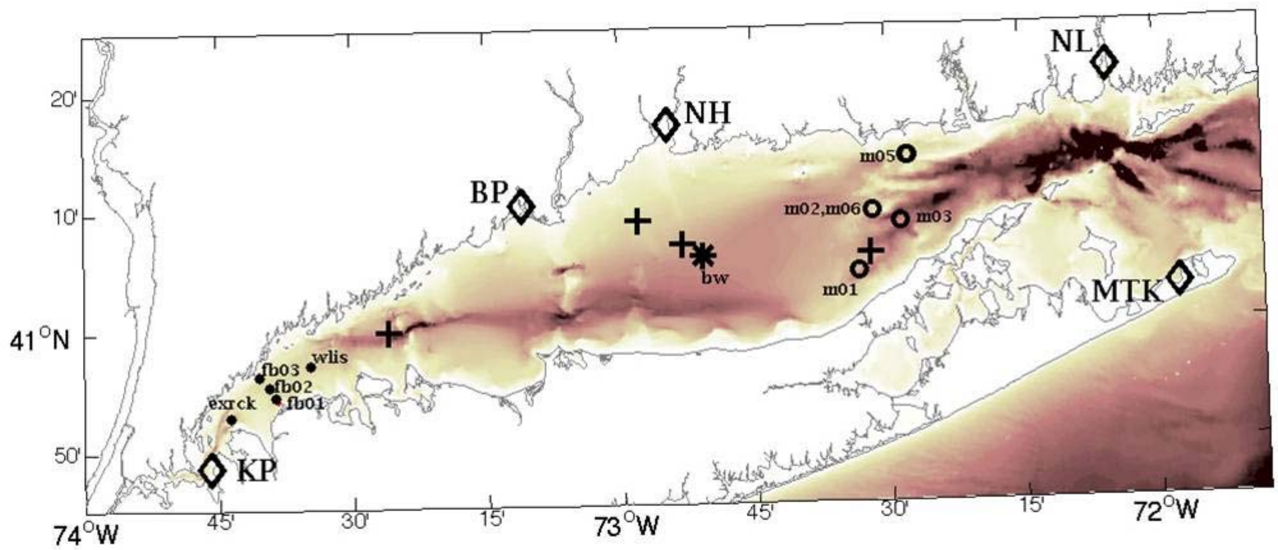


Figure 7. Bathymetric map of Long Island Sound. Diamonds show the location of NOAA tide gauges (KP = Kings Point, BP = Bridgeport, NH = New Haven, NL = New London, and MTK = Montauk), dots show LISICOS stations, asterisk shows central sound ADCP location, circles show LIS NSF stations, and pluses show historical current meter locations.

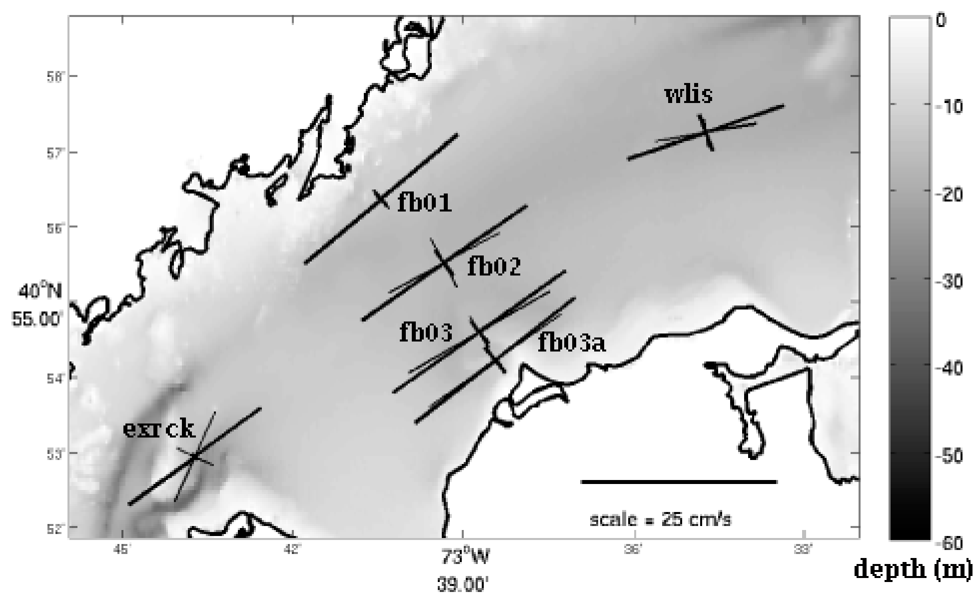


Figure 8. M2 tidal ellipses at 3m depth (heavy line) and near bottom (light line) for ADCP deployments in the western Sound. Semi-major and semi-minor axis amplitudes are shown centered at the location of the deployment.

The vertical structure of the primary tidal constituent (M2) at station fb02 is compared to the prediction of the SWEW model transport module (ECOM) in Figure 9. The results of the harmonic analysis of the M2 major and minor ellipse amplitudes are shown by the blue symbols in 9(a) and (c) for stations m02 and wlis respectively. The results from SWEM are shown by the red lines. At both stations shown, one in the eastern Sound and one in the western Sound (and the others as well), the agreement between the model and observations is excellent. The errors in the phase and its vertical structure are significant however since they affect the magnitude of the vertical shear in the horizontal currents.

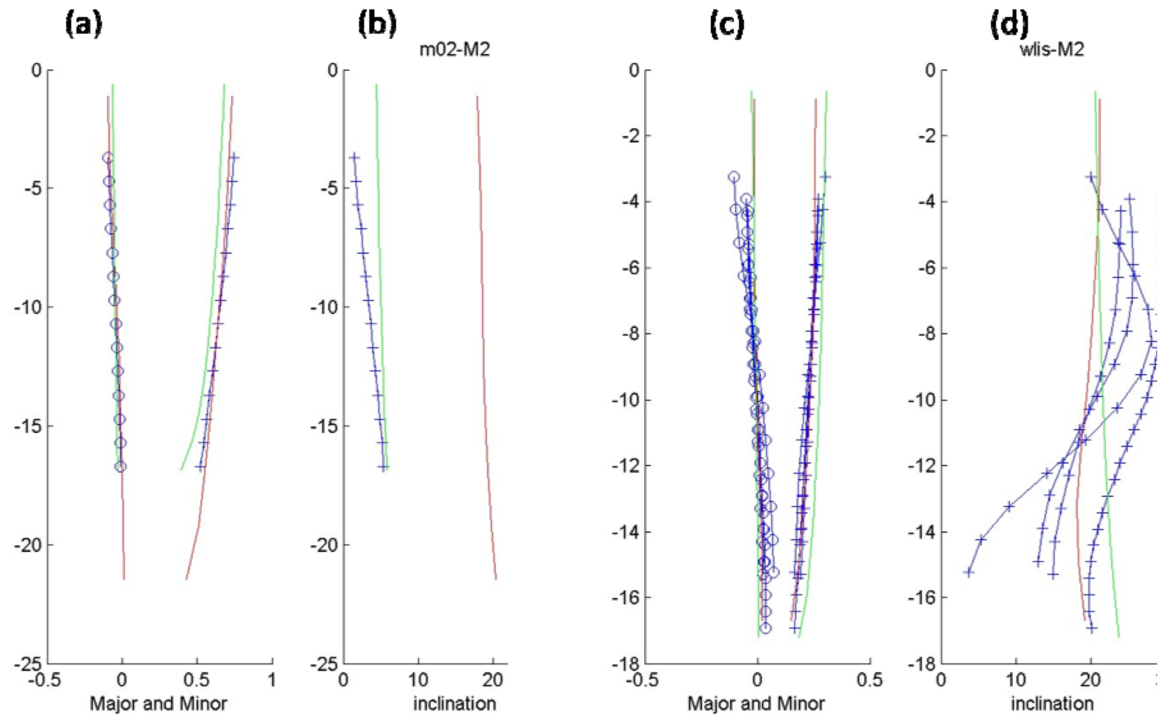


Figure 9. Comparison of the vertical structure of theM2 tidal current ellipse parameters from data (blue symbols) and the hydrodynamics module of SWEM (red lines). There are multiple profiles of observations because there were several deployments at each site. The green lines are from an alternative model that we do not discuss here. (a) and (b) show the major and minor axes (m/s) and the inclination angle relative to the along Sound direction (degrees) at station m02 in the eastern Sound. (c) and (d) show the same properties at wlis.

One of the surprising results of the analysis of Bennett et al. (2009) was the discovery of large amplitude overtides in western LIS. These are known to occur in theoretical models as a result of nonlinear dynamical processes and have been observed in some estuaries. Figure 10 shows a comparison of the predictions of SWEM for the vertical structure of the M6 ellipse axes at stations fb02 and exrk. Again the agreement of the model and the observations is very good. Note that there has been no tuning of the model at this frequency and that the developers had no knowledge of the significance of the M6 in this area. This is strong evidence that the SWEM hydrodynamics module is representing circulation well.

Though the vertical transport of DO by turbulent mixing has long been recognized as a key variable in the determination of the duration and extent of hypoxia and is simulated in SWEM, no estimates had been made from observations in LIS. In the last few years an approach has been developed to estimate the vertical eddy flux coefficients from small profiling instruments in shallow stratified shear flows and shown to yield consistent results. The SCAMP (see Ruddick, 2000), infers the rate of dissipation of turbulent kinetic energy, ϵ, from measurements of the displacement length scales, L_t, obtained from temperature microstructure measurements following the approach of Thorpe (1977). Since it has been demonstrated that the Thorpe scale, L_t, is proportional to the Ozmidov scale, $L_o = (\varepsilon / N^3)^{1/2}$ (Dillon, 1982 and Peters 1996), where $N^2 = -\frac{g}{\rho}\frac{\partial \rho}{\partial z}$ is the buoyancy or Brunt-Viasala frequency, the dissipation rate can be estimated as $\epsilon = L_t^2 N^3$. The turbulent eddy diffusivity following Osborne (1990) is $K = \frac{\epsilon}{N^2}$ so then $K = {L_T}^2 N$.

On June 27th, 2006 we took the opportunity to deploy a SCAMP microstructure profiler from the R/V Connecticut which was drifting near the exrk site shown in Figure 8. We obtained 102 vertical profiles during a 24 hour period. Since estimates of properties of turbulence are inherently noisy, we followed the advice of Thorpe (2005) and computed the 95 percentile of L_T^2values in 1m vertical intervals and 2 hour bins. We averaged N^2 in the same bins and then computed the vertical eddy coefficient for turbulent heat transfer $K = {L_T}^2 N$. The time average is shown in Figure 11. The minimum value, $3\ x\ 10^{-4}\ m^2/s$, occurs at the pycnocline. In the rest of the water column it is a factor of 10 higher. These measurements are consistent with estimates made by McCardell and O'Donnell (2009) of the seasonal evolution of 3-day mean depth-averaged effective eddy diffusivities at exrk for 2004. They found that these ranged from $1\ x\ 10^{-4}\ m^2/s$ to $1\ x\ 10^{-3}$ m^2/s.

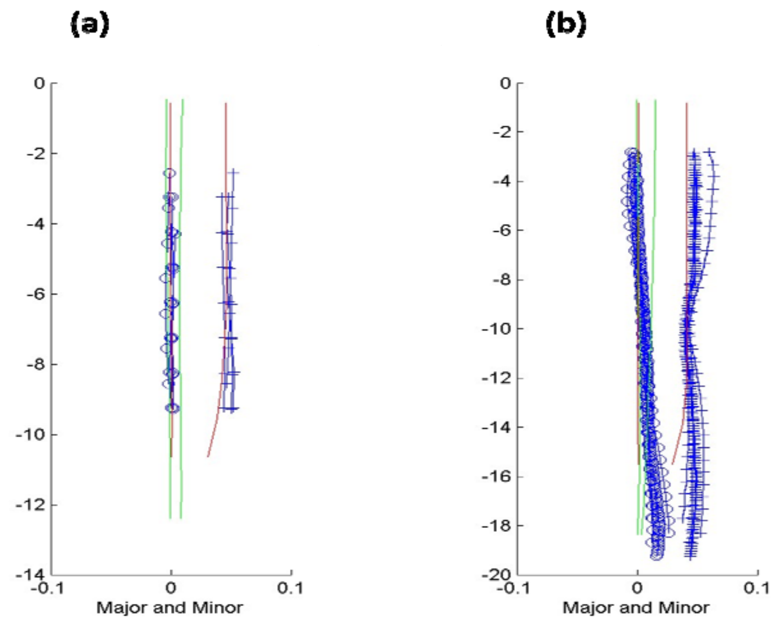


Figure 10. Comparison of the vertical structure of theM6 tidal current ellipse parameters from data (blue symbols) and the hydrodynamics module of SWEM (red lines). There are multiple profiles of observations because there were several deployments at each site. The green lines are from an alternative model that we do not discuss here. (a) and (b) show the major and minor axes (m/s) at stations fb02 and exrk.

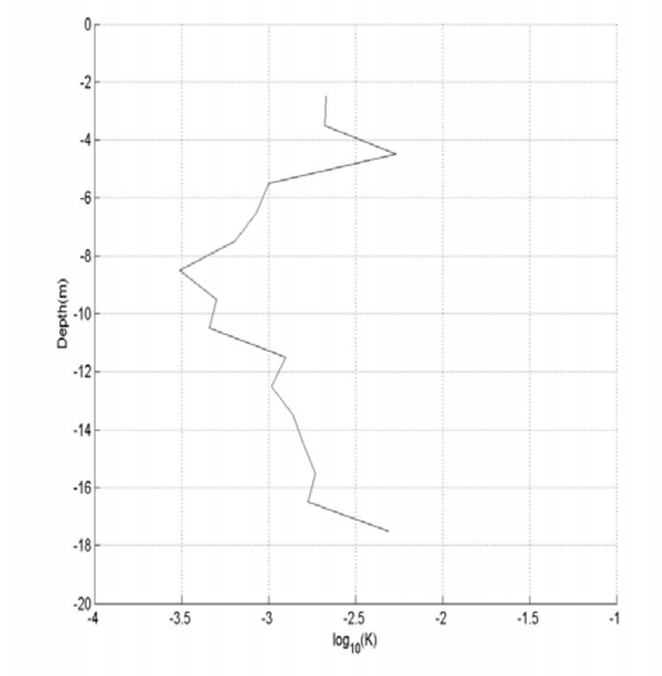


Figure 11. The time mean vertical structure of the turbulent eddy diffusivity (m^2/s) estimated using Thorpe's overturning scale from 102 profiles from a SCAMP at exrk on Jun 27th, 2006.

In ECOM, *K* is computed using the Mellor-Yamada closure (1982). We modified the SWEM code to allow output of the ECOM *K* values in order to examine how these vary during the tidal cycle in the summer. Figure 12 show the evolution of *K* at mid-depth at the model cell containing station B3, approximately 20km east of exrk. The bottom panel (12b) shows that the values produced by ECOM range from 10^{-6} to 10^{-2} m^2/s. This is not inconsistent with the values observed at exrk. The top panel (12a) shows the values implemented in the SWEM model. At the beginning of June, the ECOM magnitudes are reduced by a factor of 1/100 until September and a further reduction to 1/1000 of the original values is imposed during July. Without these reductions, the effect of the summer stratification on *K* is much more subtle.

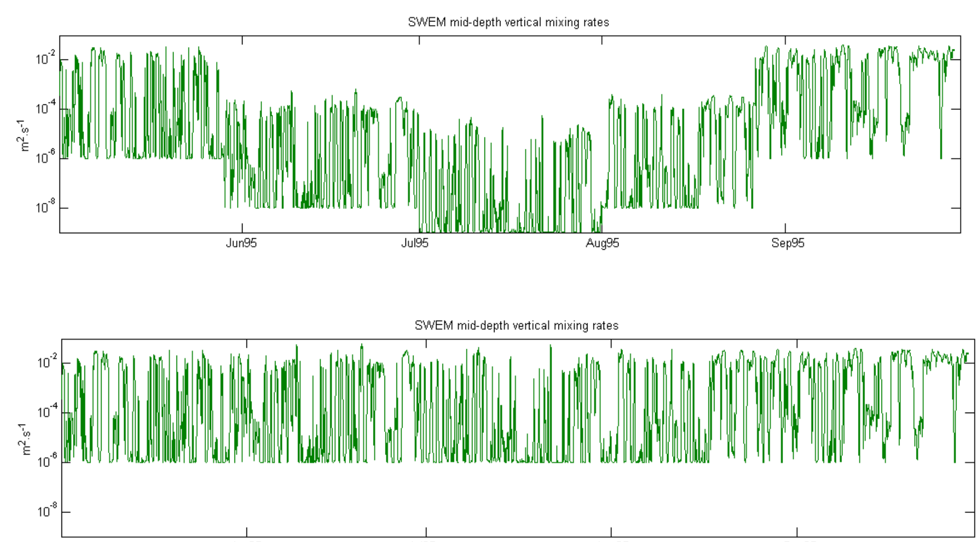


Figure 12. Comparison of the SWEM model predictions for the evolution of vertical eddy flux coefficient *K* (m^2/s) near the CTDEP station B3 in the summer of 1995. (a) displays the vertical eddy diffusion coefficient at mid-depth and (b) with the ad-hoc reduction of the vertical eddy coefficient turned off.

We conclude that the transport processes represented in SWEM by the ECOM circulation model and the Mellor-Yamada closure (prior to the ad-hoc reductions) are not inconsistent with available observations. In fact, the tidal circulation, which dominates vertical mixing variability, seems to be very well simulated, even getting the magnitudes of the higher tidal harmonics correctly. However, the ad-hoc summertime reductions to these values are unrealistic and are inconsistent with available observations.

7.4 Task 4. Collate Data and Maintain an Archive.

We have acquired the data obtained by the CTDEP Long Island Sound water quality survey program between 1991 and 2009 and archived it in a Microsoft SQLserver database. This data set includes the properties listed in Table 3 at the stations shown in Figure 13.The database management software has also been configured to archive the real-time data from the LISICOS buoy array and CODAR network. We also have implemented ArcView to generate visual representations of available data and the ArcIMS mapserver tool to view them over the internet. A forms-based page allows users to select and download data from the archive in as ASCII or Excel files. Time series at a station, or all data from a month can be downloaded. The site can be accessed at www.LISICOS.UConn.edu.

Table 3. Water quality variables in the evaluation database of observations acquired by the CT DEP.

Variable	Symbol	Unit
Biogenic Silica	BIOSI-LC	mg/L
Chlorophyll a (<0.7 microns)	CHLA	mg/L
Chlorophyll a (<0.5microns)	CHLA0.5	mg/L
Orthophosphate	DIP	mg/L
Dissolved Organic Carbon	DOC	mg/L
Ammonia	NH#-LC	mg/L
Nitrate + Nitrite	NOX-LC	mg/L
Particulate Carbon	PC	mg/L
Particulate Nitrogen	PN	mg/L
Particulate Phosphorus	PP-LC	mg/L
Dissolved Silica	SIO2-LC	mg/L
Total Dissolved Nitrogen	TDN-LC	mg/L
Total Dissolved Phosphorus	TDP	mg/L
Total Suspended Solids	TSS	mg/L

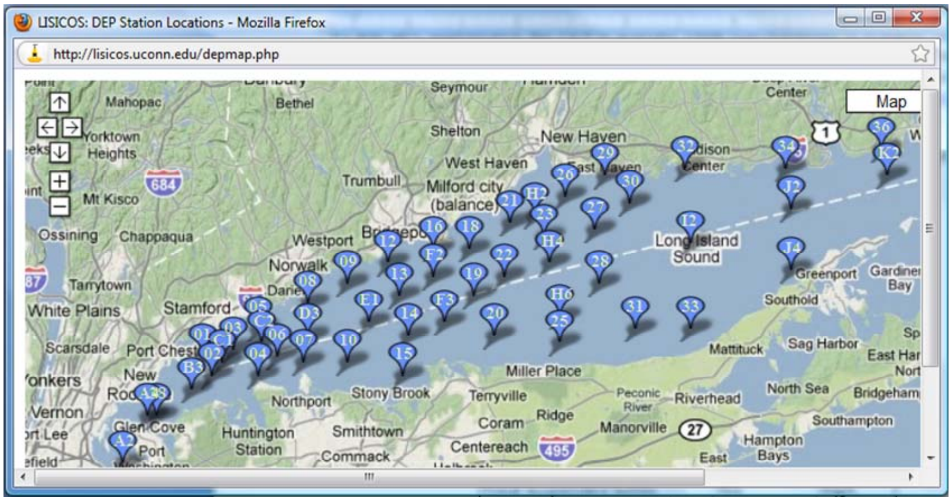


Figure 13. Station locations used in the database.

7.5. Task 5. Sensitivity of the Choices of Model Parameters.

We evaluated the sensitivity of the model to the choice of parameters using the model skill defined in Equation (1) and the monthly mean reference model. The location of the data stations used are shown in Figure 14. At this subset of the stations shown in Figure 13, measurements are available for both evaluation years (1989 and 1995). Figure 15 shows the full SWEM computational domain as a square matrix with the ocean shaded blue and the land shaded green. The model grid points used in the evaluation are shaded red. It is important to note that the elements of the model solution examined in Long Island Sound form a small subset of the full SWEM domain and our assessment of skill applies to that small area of the model.

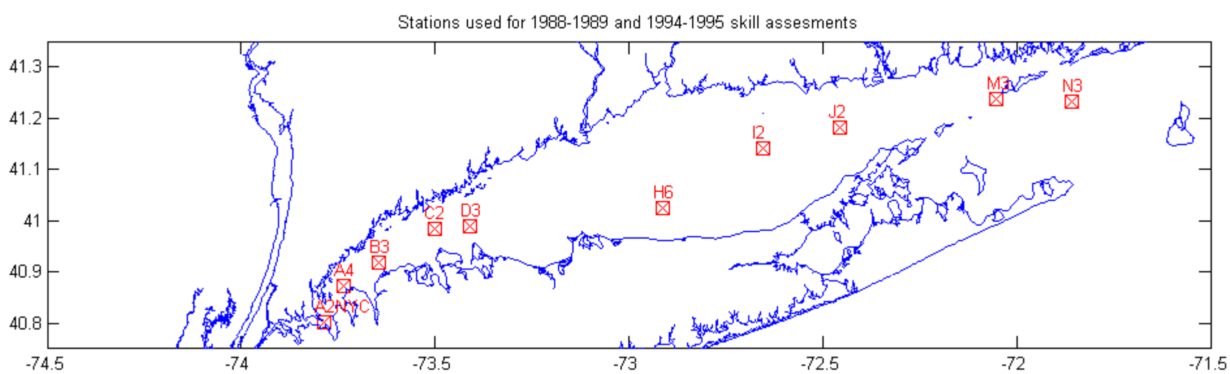


Figure 14. Coastline of Long Island Sound together with the locations of the CT DEP survey locations used in the model evaluation.

.

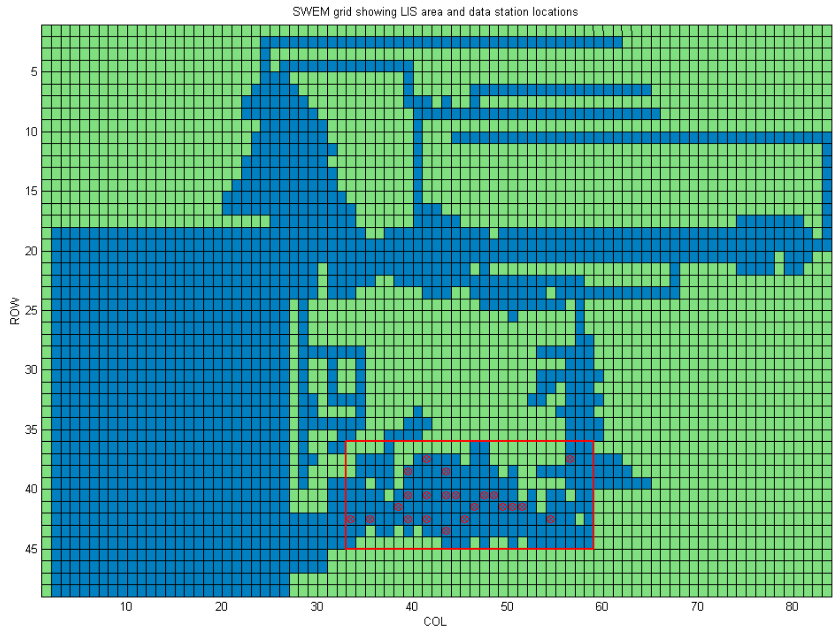


Figure 15. The SWEM domain with the land shaded green and the ocean shaded blue. The cells that simulate LIS are within the red box.

The SWEM model contains 120 parameters that were used by Hydroqual to "tune" SWEM. Many values were transferred from other model applications and are thought to be fairly robust. Others are uncertain and are, largely, empirical coefficients representing unresolved or poorly understood processes. A complete list of these parameters and a brief reference to their scientific meanings is listed in Table 4. The mathematical formulations are defined in Hydroqual (2003).

As a primary assessment of the sensitivity of the predictions to the parameter choices we perturbed each parameter value from that used by Hydroqual (2004) by ±10% and recomputed the SWEM solutions, the overall skill, and the skill for each of the 12 model variables for which data was available (See Table 2). This required a total of 240 model runs for each model year. Figure 16 shows the changes in the overall skill associated with the perturbation of certain parameters. The basic-case skill is shown by the horizontal lines. The results are also summarized in the first column of Table 4 where the change in skill per unit change in the parameter values is listed. The parameters that showed the most influence on S_1 were identified by ranking them by $\left|\frac{\Delta S}{\Delta P}\right|$ (see Table 4). Not surprisingly, the top 10 were all associated with the temperature dependence of rate coefficients.

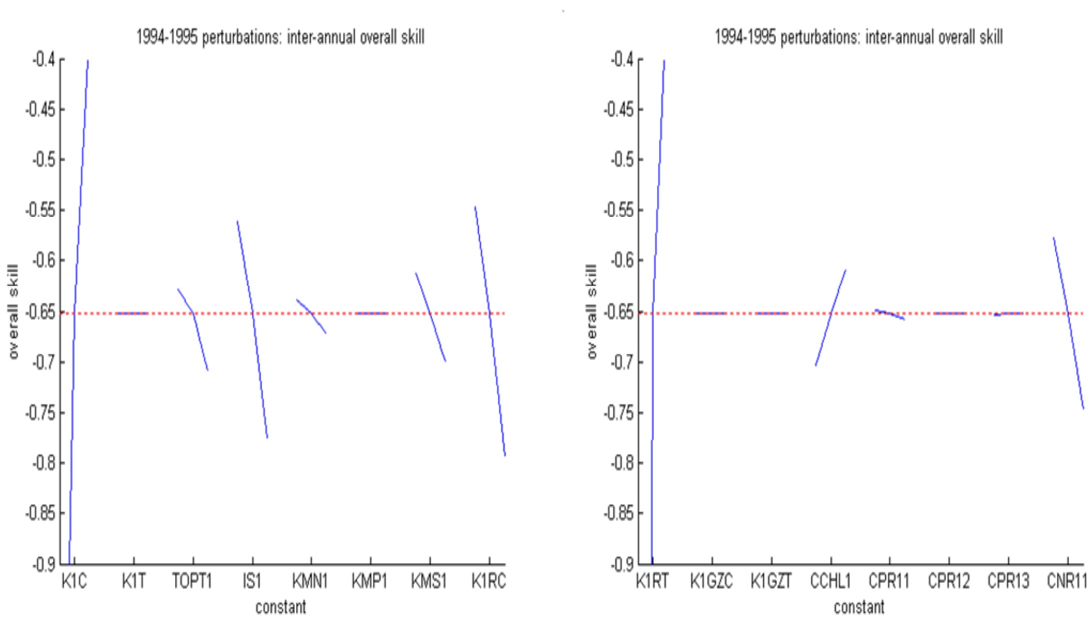


Figure 16. Change in Skill, S_1, for the 1995 simulation resulting from changes to 16 model parameters and the 1995 simulation. The effect of increasing the parameter value by 10% is shown on the right side of each line segment and the effect of decreasing it is shown on the left.

We then computed solutions for eight different values of each of 8 parameters for which the skill showed high sensitivity. The values chosen were evenly spaced throughout the range suggested by Hydroqual as reproduced in Table 4. Figure 17 shows the overall skill dependence on the four parameters K1C, K1RC, K1RT, and CCHL1. The values Hydroqual chose for their simulations are shown by circles and the vertical lines. As discussed above, K1C controls the spring plankton community growth rate. K1RC controls the rate of respiration by the spring phytoplankton and the removal of plankton biomass. K1RT describes the influence of temperature on the respiration rate of the spring plankton and the removal of plankton biomass. The fourth parameter, CCHL1, is the carbon (the modeled quantity) to chlorophyll ratio. It does not influence the model dynamics but is critical to the comparison of model results to observations of chlorophyll.

It is clear that both the 1989 and 1995 simulations could be substantially improved by modification of parameter choices. Increasing K1C by 50%, for example, would cause the skill of the 1989 simulation to be positive. Presumably this would increase the plankton biomass. It is also evident that coefficients that influence the spring respiration (K1RC and K1RT) are already at their near optimal values within the range that Hydroqual considered scientifically plausible. However, these parameters are likely to be coupled in a nonlinear manner and an extensive stepwise approach to improving the calibration would be necessary. This is not within the scope of the present project and, as we will show, more substantial changes are necessary first.

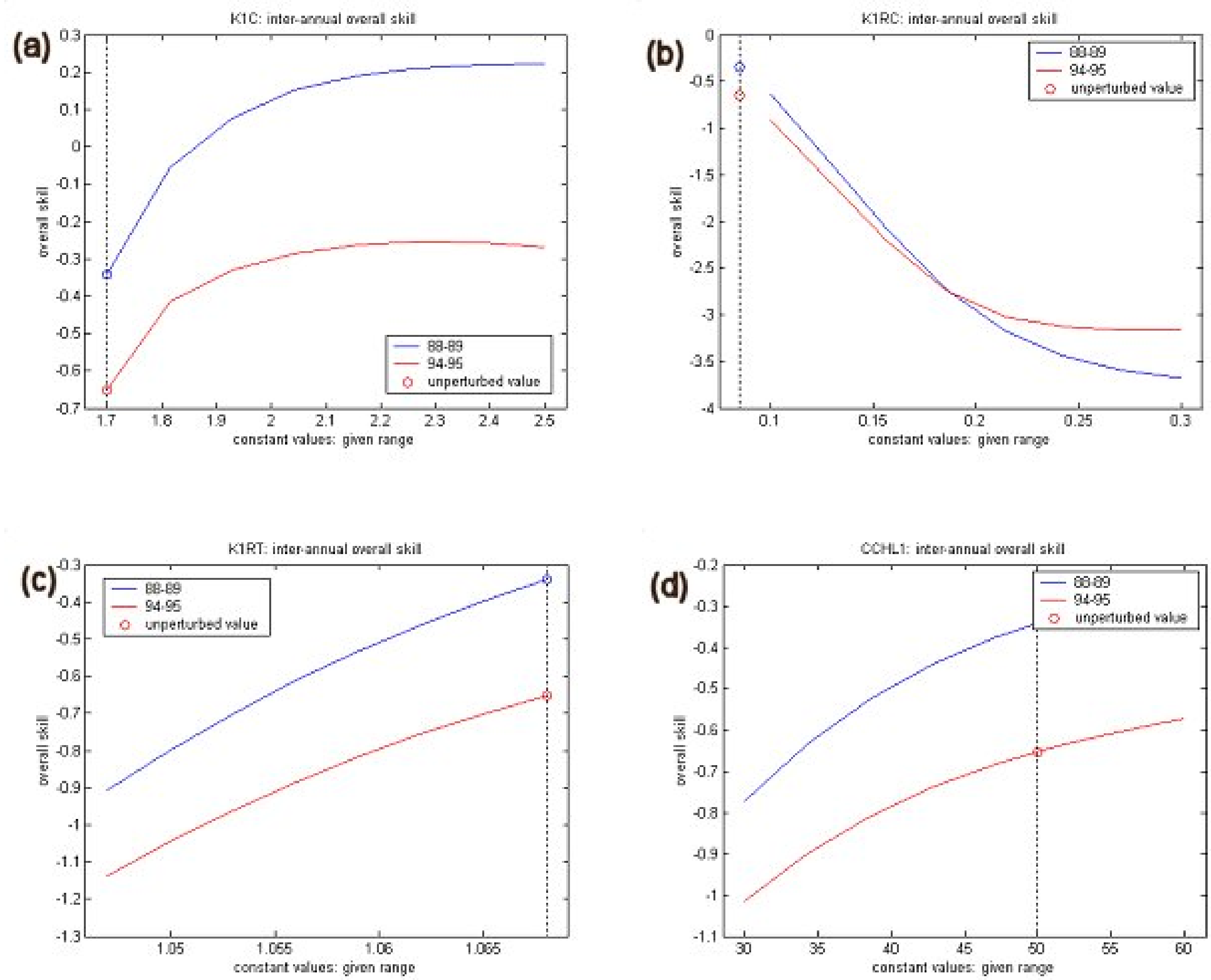


Figure 17. The dependence of the Skill, S_1, for the 1989 (blue) and 1995 (red) simulations on the choice of the values for (a) K1C, (b) K1RC, (c) K1RT, and (d) CCHL1.

Table 4. Parameters used in SWEM.

Sensitivity d/dC(Skill)	NO	NAME	DESCRIPTION	VALUES	UNITS	REFERENCE VALUE
15.54	9	K1RT	TEMPERATURE COEFFICIENT	1.068		1.047 - 1.068
5.74	104	VSBAST	TEMPERATURE CORRECTION	1.029		1.029
5.55	1	K1C	SATURATED PHYTOPLANKTON GROWTH RATE (AT TEMPERATURE = TOPT1)	1.7	/DAY	1.7 - 2.5
4.79	110	VSSEDT	TEMPERATURE CORRECTION FOR DEPOSITION TO SEDIMENT	1.029		
4.72	72	K1113T	TEMPERATURE COEFFICIENT	1.08		1.08
2.23	82	K1516T	TEMPERATURE COEFFICIENT	1.08		1.08
1.55	95	FLOCEX	FRACTION OF PRIMARY PRODUCTIVITY GOING TO LABILE ORGANIC CARBON VIA EXUDATION	0.2		
1.54	70	K1012T	TEMPERATURE COEFFICIENT	1.08		1.08
1.51	74	K1213T	TEMPERATURE COEFFICIENT	1.08		1.08
1.47	107	VSPMT	TEMPERATURE CORRECTION	1.029		
1.30	8	K1RC	ENDOGENOUS RESPIRATION RATE AT 20 DEG C	0.085	/DAY	0.1 - 0.3
1.26	44	KMPHYT	HALF SATURATION CONSTANT FOR PHYTOPLANKTON	0.0	MG C/L	0.05
1.13	4	IS1	SATURATING ALGAL LIGHT INTENSITY	150.0	LY/DAY	
1.02	24	TOPT2	OPTIMAL GROWTH TEMPERATURE FOR SUMMER ASSEMBLAGE	26.0	DEG C	20.0 - 25.0
0.95	101	KAT	TEMPERATURE CORRECTION COEFFICIENT FOR ATMOSPHERIC REAERATION	1.024		
0.89	76	K1314T	TEMPERATURE COEFFICIENT	1.08		1.08
0.89	16	CRBN11	CARBON TO NITROGEN RATIO - NON-N LIMITED	5.67	MG C/MG N	5.2 -5.67
0.80	54	FNH3	AMMONIA	0.35		0.15 - 0.35
0.62	22	K2C	SATURATED PHYTOPLANKTON GROWTH RATE (AT TEMPERATURE = TOPT2)	3.0	/DAY	2.0 - 3.0
0.54	30	K2RT	TEMPERATURE COEFFICIENT	1.068		1.047 - 1.068
0.50	43	XKC	CHLOROPHYLL SELF-SHADING EXT.COEFF.	0.017	M2/MG CHLA	0.017
0.50	12	CCHL1	CARBON TO CHLOROPHYLL RATIO	50.0	MG C/MG CHLA	30.0 - 60.0
0.46	7	KMS1	HALF SATURATION CONSTANT FOR SILICA	0.02	MG SI/L	0.02
0.45	17	CRBN12	CARBON TO NITROGEN RATIO - N LIMITED	7.2	MG C/MG N	6.5 - 7.2
0.42	3	TOPT1	OPTIMAL GROWTH TEMPERATURE FOR DIATOMS	8.0	DEG C	6.0 - 12.0
0.40	19	CRBS11	CARBON TO SILICA RATIO - NON-SI LIMITED	2.5	MG C/MG SI	2.2 - 3.0
0.39	71	K1113C	(MINERAL) MINERALIZATION RATE OF RDON TO NH3	0.008	/DAY	0.008 - 0.01
0.32	88	K190T	TEMPERATURE COEFFICIENT	1.08		1.08
0.29	20	CRBS12	CARBON TO SILICA RATIO - SI LIMITED	8.0	MG C/MG SI	8.0 - 15.0
0.26	21	CRBS13	COEFFICIENT DETERMINING RANGE OF SI LIMITATION	12.0	L/MG SI	12.0 - 30.0
0.26	81	K1516C	MINERALIZATION RATE OF BIOGENIC SI TO AVAIL SI	0.08	/DAY	0.1 - 0.25
0.24	64	K68T	TEMPERATURE COEFFICIENT	1.08		1.08
0.22	102	VSBASE	BASE ALGAL SETTLING RATE	0.2	M/DAY	0.2 - 1.0
0.21	51	FLPON	LABILE PARTICULATE ORGANIC NITROGEN	0.3		0.3 - 0.35
0.21	53	FLDON	LABILE DISSOLVED ORGANIC NITROGEN	0.125		0.125 - 0.15
0.17	5	KMN1	HALF SATURATION CONSTANT FOR NITROGEN	0.01	MG N/L	0.01
0.16	66	K78T	TEMPERATURE COEFFICIENT	1.08		1.08
0.16	29	K2RC	ENDOGENOUS RESPIRATION RATE AT 20 DEG C	0.125	/DAY	0.10 - 0.30
0.15	69	K1012C	HYDROLYSIS RATE OF LPON TO LDON	0.05	/DAY	0.05 - 0.07
0.15	79	K140T	TEMPERATURE COEFFICIENT	1.045		1.045
0.14	73	K1213C	MINERALIZATION RATE OF LDON TO NH3	0.085	/DAY	0.085 - 0.1
0.12	105	VSPOM	PARTICULATE ORGANIC MATTER SETTLING RATE	1.0	M/DAY	
0.11	75	K1314C	NITRIFICATION RATE AT 20 DEG C	0.05	/DAY	0.05 - 0.1
0.10	18	CRBN13	COEFFICIENT DETERMINING RANGE OF N LIMITATION	10.0	L/MG N	10.0 - 15.0
0.10	68	K911T	TEMPERATURE COEFFICIENT	1.08		1.08
0.09	25	IS2	SATURATING ALGAL LIGHT INTENSITY	350.0	LY/DAY	
0.09	86	K1820T	TEMPERATURE COEFFICIENT	1.08		1.08
0.07	60	K46T	TEMPERATURE COEFFICIENT	1.08		1.08
0.07	90	K200T	TEMPERATURE COEFFICIENT	1.08		1.08
0.06	26	KMN2	HALF SATURATION CONSTANT FOR NITROGEN	0.01	MG N/L	0.01
0.06	40	CRBS21	CARBON TO SILICA RATIO - NON-SI LIMITED	10.0	MG C/MG SI	5.0 - 10.0
0.06	62	K57T	TEMPERATURE COEFFICIENT	1.08		1.08
0.05	97	K220T	TEMPERATURE COEFFICIENT	1.08		1.08
0.05	52	FRDON	REFRACTORY DISSOLVED ORGANIC NITROGEN	0.125		0.125 - 0.2
0.05	103	VSNUTR	NUTRIENT STRESSED ALGAL SETTLING RATE	0.5	M/DAY	0.5 - 1.0
0.04	13	CRBP11	CARBON TO PHOSPHORUS RATIO - NON-P LIMITED	40.0	MG C/MG P	25.0 - 40.0
0.04	37	CRBN21	CARBON TO NITROGEN RATIO - NON-N LIMITED	5.67	MG C/MG N	4.0 - 5.67
0.04	87	K190C	OXIDATION RATE OF RDOC	0.008	/DAY	0.007 - 0.01
0.03	47	FRDOP	REFRACTORY DISSOLVED ORGANIC PHOSPHOROUS	0.1		0.1 - 0.15
0.03	56	FLPOC	LABILE PARTICULATE ORGANIC CARBON	0.4		0.3 - 0.4
0.03	91	KMLDOC	MICHAELIS CONSTANT FOR LDOC	0.0	MG C/L	0.1
0.03	38	CRBN22	CARBON TO NITROGEN RATIO - N LIMITED	10.0	MG C/MG N	7.5 - 10.0
0.03	33	CCHL2	CARBON TO CHLOROPHYLL RATIO	100.0	MG C/MG CHLA	75.0 - 100.0
0.03	49	FPO4	DISSOLVED INORGANIC PHOSPHOROUS	0.45		0.2 - 0.45
0.03	63	K68C	MINERALIZATION RATE OF RDOP TO PO4	0.01	/DAY	0.01 - 0.02
0.02	46	FLPOP	LABILE PARTICULATE ORGANIC PHOSPHOROUS	0.25		0.25 - 0.35
0.02	58	FLDOC	LABILE DISSOLVED ORGANIC CARBON	0.45		0.35 - 0.45
0.02	65	K78C	MINERALIZATION RATE OF LDOP TO PO4	0.1	/DAY	0.1 - 0.2
0.02	84	K1719T	TEMPERATURE COEFFICIENT	1.08		1.08

7.6 Task 6. Sensitivity of Predictions to Model Formulation

By the formulation of a water quality model we mean the choice of the equations and boundary conditions. These prescribe how the fields (concentrations of nutrients, plankton etc.) evolve through addition and subtraction by sources and sinks and through interactions. In this section we report the results of our examination of the sensitivity of the predictions to boundary fluxes. This analysis led us to investigate the choice of equations and the discovery of ways to improve the model skill.

7.6.1 Boundary Flux Sensitivity

The SWEM model requires initial conditions (IC) and boundary conditions (BC) for all variables, and forcing conditions (FC) that describe discharges into the interior of the model domain. Initial conditions are read in from a file created by the hydrodynamic model and surveys. The model is then run for 5-7 yearly cycles with the model state at the end of each cycle used as the ICs for the next cycle. This allows for the consequences of inconsistencies in the initial conditions to be damped as well as allowing the benthic submodel to stabilize.

At the 74 BC cells identified in Figure 18a, Dirichlet conditions are enforced: i.e. the values of the 25 variables are specified. The loading inputs are imposed as Neumann conditions at the cells identified in Figure 1b-d: i.e. fluxes of the 25 variables are specified. These loadings are further broken down as point source (PS), non point source (NPS), fall line (FL), and atmospheric (ATM) loadings. The PS loads are added into the bottom two model layers, whereas both the FL and the NPS loads are added into the top two layers. The ATM loadings are deposited into the surface layer of **all** SWEM cells equally. Figure 18a shows the locations of the BC cells along with the PS (b), NPS (d), and FL (c) cells. Note that the locations (and layers) of the FL and NPS cells are the same.

There are significant differences between the loadings used for 88-89 and 94-95, not only in amounts, but in what is used. The field program provided additional information for 94-95 and this allowed the separation of the phosphorus and nitrogen loads into reactive and labile components. The PS loads for 88-89 are constant for the year, whereas for 94-95 these vary by month. The NPS and FL loads vary by day for both years but some systems have no loadings for 88-89 but do have loadings for 94-95. The ATM loadings vary by month for both years but some systems have no loadings for 88-89 but do have loadings for 94-95. Because of the disparities in how these loadings are represented, a direct comparison of these is best done by looking at the mean daily loadings integrated across the SWEM spatial domain. Table 5 summarizes these results with the substantial changes shaded for emphasis. It is not clear that the additional forcing detail available in 1994-95 led to a better simulation of DO.

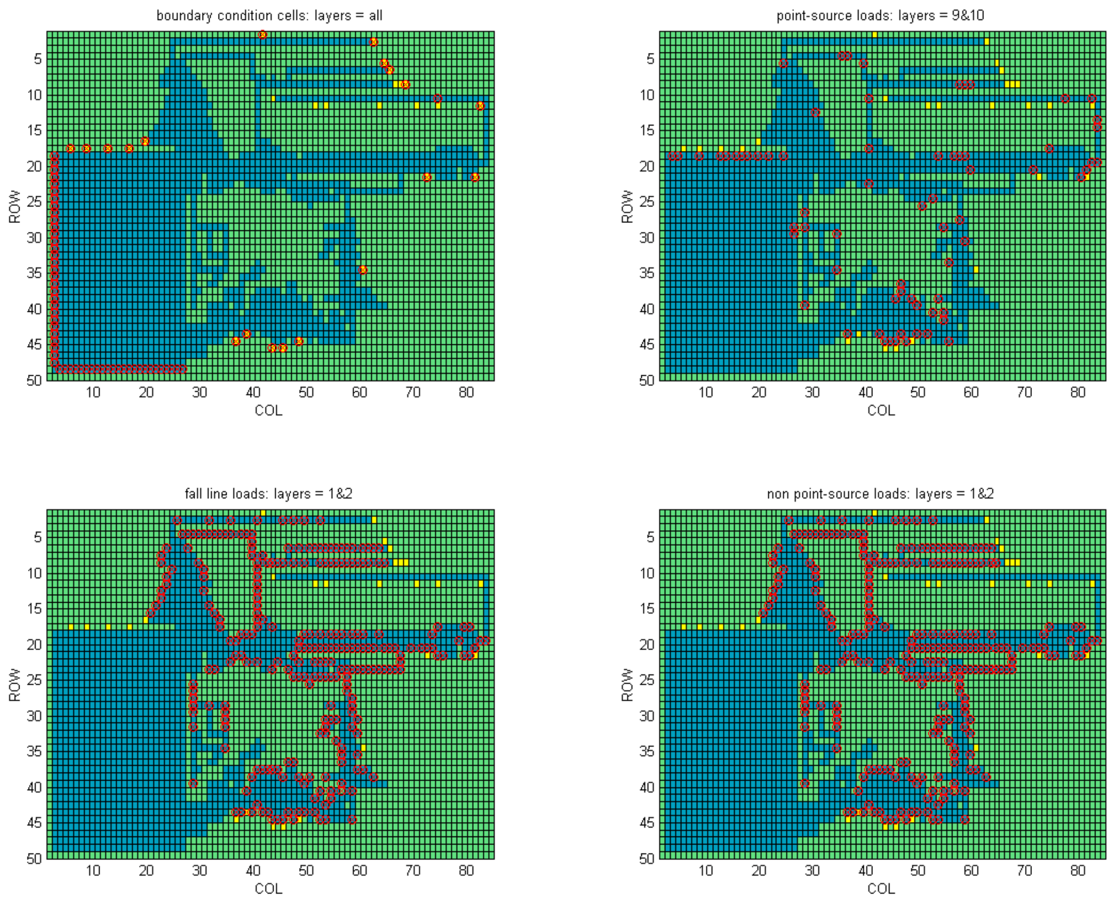


Figure 18. Grid locations of BC (Panel A), PS (Panel B), FL (Panel C), and NPS (Panel D) cells.

Our experiments with SWEM showed that the solutions were relatively insensitive to the magnitudes of the PS and NPS discharges prescribed. To demonstrate this we summarize the results of simulations with large, ±100%, changes in the loadings in a simulation of the 1994-1995 period. This effectively simulates the impact of eliminating and doubling the loadings.

In Figure 19 we show the predicted [O2] at three levels (surface, mid depth and bottom) at the cell containing CTDEP station C1 which is in the area usually subject to hypoxia. The blue line shows the solution for the standard (base) case and the green and red lines show the solutions when the PS and NPS loads are set to zero and doubled respectively. At the surface (layer 1) there is little difference in [O2] until mid-June when the green line is significantly lower, ~2 mg/l, than the other two cases. This difference between the green line and the others can be interpreted as a measure of the stimulation of production by the nutrients from the PS and NPS fluxes. There is no significant difference in the predictions at mid-depth (layer 5). At the bottom (layer 10) the green line is slightly, ~0.5 mg/l, above the red and blues lines from July to mid-August. These differences suggest doubling the loading will not appreciably decrease the minimum [DO] at C1 and removing them will increase the minimum [DO] by less than 0.5 mg/l. The stations to the east show even less response to changes in loading perturbations. We conclude that SWEM appears to be very insensitive to even huge changes in the PS and NPS loadings.

Table 2. Total mean daily loads for 1988-1989 and 1994-1995 SWEM runs (metric tons per day). Substantial changes between years have been emphasized by the color shading.

System	PS 88-89	PS 94-95	NPS 88-89	NPS 94-95	FL 88-89	FL 94-95	ATM 88-89	ATM 94-95	Total 88-89	Total 94-95
SAL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PHYT1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PHYT2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RPOP	0.00	1.06	0.00	0.30	0.00	0.28	0.00	0.00	0.00	1.64
LPOP	3.53	2.48	1.71	0.13	1.28	0.28	0.00	0.00	6.52	2.88
RDOP	0.00	0.73	0.00	0.06	0.00	0.05	0.00	0.00	0.00	0.84
LDOP	3.53	1.69	0.16	0.03	0.09	0.05	2.01	1.07	5.79	2.85
DIP	14.93	18.31	0.62	0.40	0.52	0.48	2.45	0.86	18.53	20.05
RPON	0.00	12.12	0.00	1.24	0.00	1.22	0.00	0.00	0.00	14.57
LPON	35.85	28.28	1.09	0.53	1.98	1.22	0.00	0.00	38.92	30.03
RDON	0.00	2.55	0.00	1.34	0.00	0.65	0.00	0.00	0.00	4.55
LDON	35.25	5.95	3.75	0.57	2.72	0.65	24.52	15.41	66.23	22.59
NH3	112.22	148.31	2.07	1.12	2.71	3.57	38.44	42.68	155.43	195.68
NO23	14.50	17.67	3.75	3.63	1.03	0.40	132.49	99.04	151.77	120.73
SIU	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SI	39.59	28.17	8.45	8.38	2.47	1.38	0.00	3.85	50.51	41.77
RPOC	46.10	28.41	0.00	24.30	0.00	16.69	0.00	0.00	46.10	69.39
LPOC	69.15	66.27	38.95	10.41	67.06	16.69	0.00	0.00	175.17	93.37
RDOC	80.56	36.53	0.00	29.24	0.00	7.53	0.00	0.00	80.56	73.29
LDOC	80.56	36.53	15.70	6.26	23.14	3.76	144.86	108.70	264.25	155.26
REDOC	161.10	73.06	0.00	6.26	0.00	3.76	144.86	108.70	305.96	191.79
EXDOC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
O2EQ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DO	65.60	60.35	24.41	30.03	6.00	3.07	0.00	0.00	96.01	93.44
TEM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

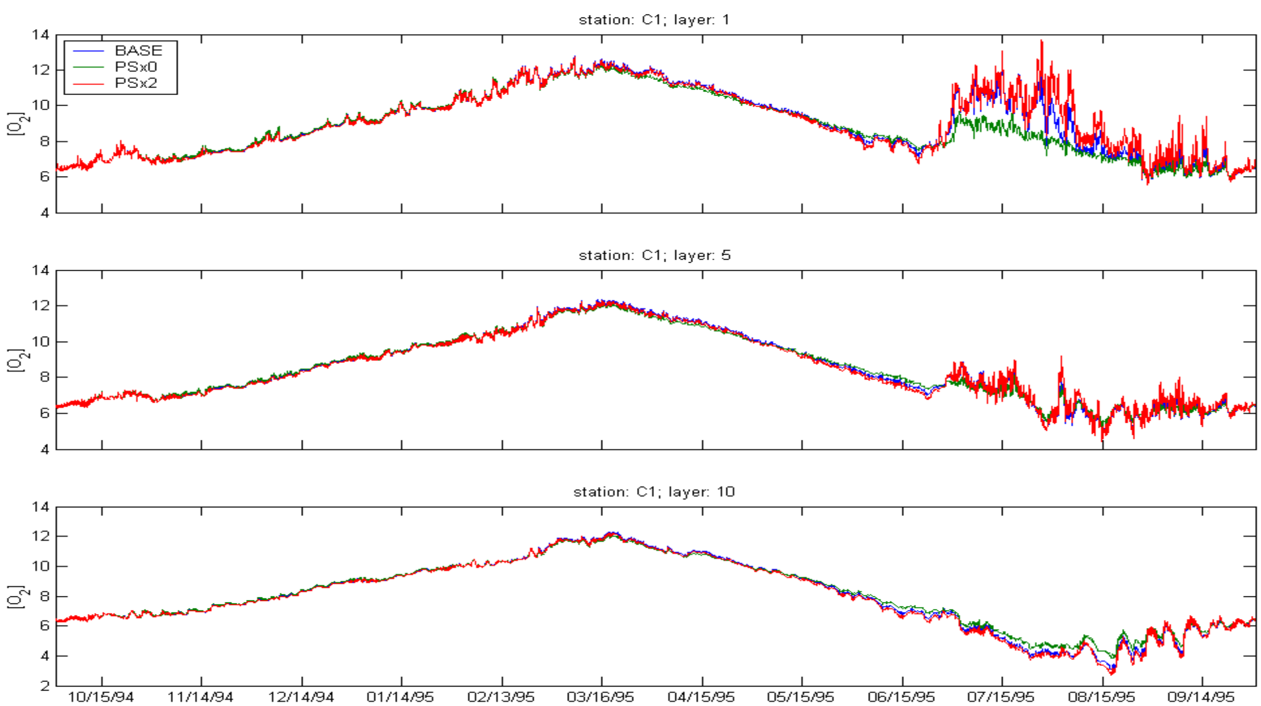


Figure 19. SWEM PS loading perturbations at station C1 (cell 41,50); blue: unchanged, green: no PS loading; red: doubled PS loading; top panel: surface layer; middle panel: mid-depth layer; bottom panel: bottom layer.

To understand why changing the PS nutrient loadings has such a minor effect on the SWEM predictions requires examining the phytoplankton dynamics. SWEM includes two algal populations: PHYT1 (system 2) represents winter diatoms and PHYT2 (system 3) summer flagellates. SWEM represents growth in PHYT1 and PHYT2 using the conventional nutrient limitation factor which is calculated as:

$$nutrient\,limitation = \min\left(\frac{[\mathrm{Si}]}{[\mathrm{Si}]+K_{\mathrm{Si}}}, \frac{[\mathrm{NH_3}]+[\mathrm{NO_{2,3}}]}{[\mathrm{NH_3}]+[\mathrm{NO_{2,3}}]+K_{\mathrm{N}}}, \frac{[\mathrm{PO_4}]}{[\mathrm{PO_4}]+K_{\mathrm{P}}} \right) \qquad (2)$$

where $K_{\mathrm{Si,N,P}}$ are half-saturation constants for silica, nitrogen, and phosphorus.

Examination of the variation of the three ratios in Equation (2) computed from the solution shows that nutrient limitation is less important for the PHYT1 spring assemblage but plays a major role in for the PHYT2 summer assemblage, beginning in late June. This explains why the model showed more sensitivity to the temperature dependent maximum growth parameters for PHYT1 than it did for PHYT2 and also why perturbing the loadings had little effect on PHYT1.

Further, it appears that the model shows little sensitivity to the PS and NPS loading perturbations because the productivity and growth of the summertime PHYT2 population (which in reality is largely dinoflagellates) is nutrient limited in the model. Note that Equation (2) includes the silica concentration in the computation of both winter diatom and summer flagellate population dynamics: the formulations used for the two populations differed only in the values of the half-saturation constants.

Figure 20a shows the June-September, 1995 variation of the SWEM nutrient limitation factor, Equation (2), for the surface layer near C1 and (b-d) show the evolution of the ammonia+nitrate, phosphate, and silica levels in the model at this location. The blue lines show the original SWEM results and the green and red lines show what happens when the PS and NPS conditions are set to zero and doubled. The dashed lines in each panel show the half concentration constants for PHYT2 for these three nutrients (SWEM constants KMN2, KMP2 and KMS2). In Figure 20a, the red line shows that doubling the point sources causes the nutrient limitation factor to rapidly fluctuate between 0 and 1. In contrast, the green line shows modest nutrient limitation throughout the summer. Figure 20d shows that in both the standard (blue lines) and the increased loading (red lines) simulations the Si concentration becomes very small in early July and this appears to limit further growth. Reducing the point source loadings to zero reduces the productivity and uptake of Si and PO4 leaving TDN to be the growth limiting factor. When the point source loadings are increased, however, the model switches from being nitrogen limited to being silica limited. This may be a model artifact caused by the use of the same formulation for nutrient limitation for both summer and winter populations. The half saturation value for Si used for the summer formulation is greatly reduced from that used for the winter population, but this does not prevent silica from becoming a limiting nutrient when there is sufficient nitrogen and phosphorus. This aspect of the SWEM dynamics requires further investigation as there are no observations or studies of which we are aware that show silica limitation to be a controlling factor for LIS summertime algal dynamics.

7.6.2 Mixing Rate Sensitivity

We noted in Section 7.3 that SWEM as implemented by Hydroqual required a substantial reduction in the summertime vertical mixing rates in the WLIS to achieve realistic simulations of the summertime oxygen depletion in the near bottom waters. Recent observational evidence (O'Donnell et al, 2008; McCardell and O'Donnell, 2009) suggests that the vertical transport of oxygen by vertical mixing may be responsible for a significant flux and much of the variability in the near bottom oxygen levels, so the elimination of all vertical transport as is currently implemented in the model seems unwise. New observations of the rate of primary production and respiration in Long Island Sound have become available in the course of the project (Goebel and Kremer, 2007) and observations of the evolution of near surface DO variability were published by O'Donnell et al. (2008) which can provide a guide to the magnitude of near the surface production rate (Odum, 1956). An evaluation of the DO budget in SWEM is, therefore, possible.

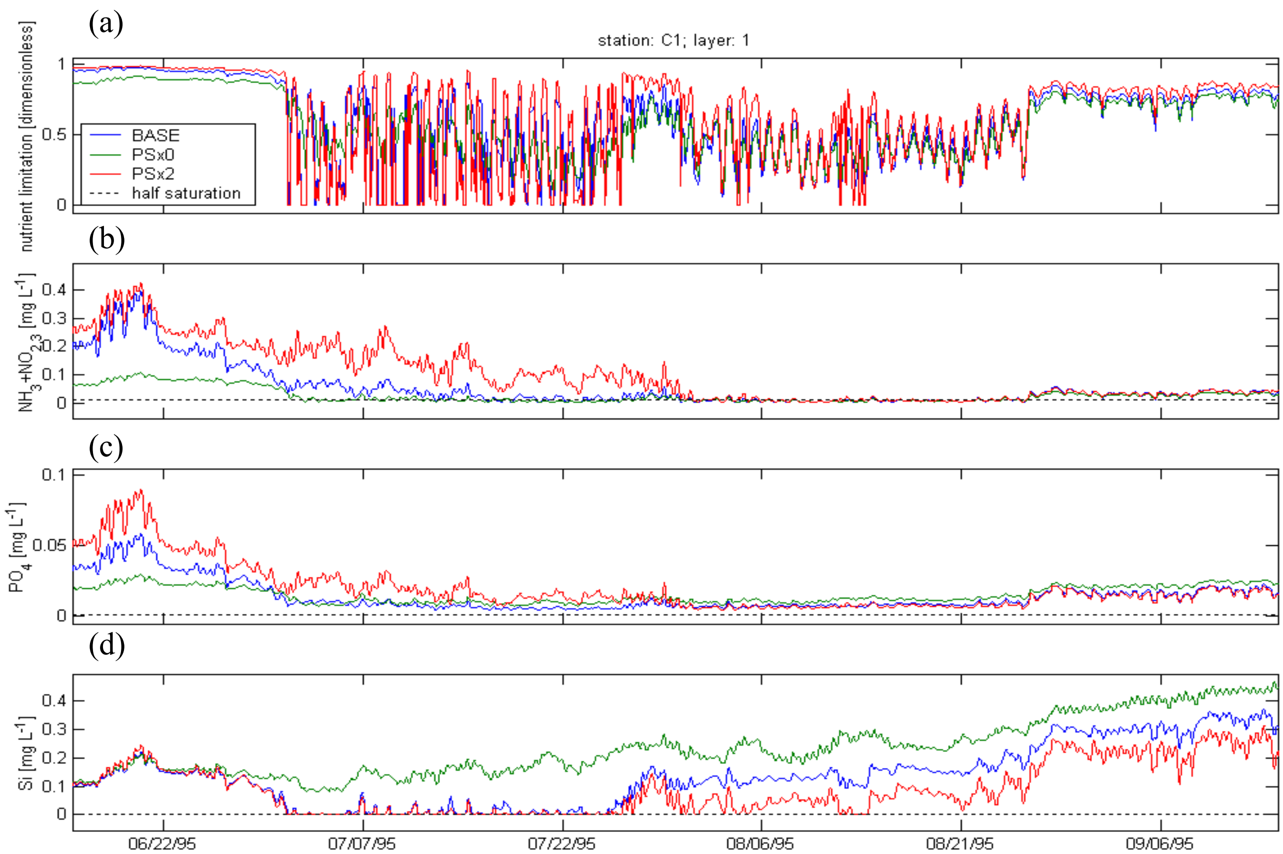


Figure 20. June-September nutrient limitation for PHYT2 and nutrient levels with PS loading perturbations at station C1 (cell 41,50); blue: unchanged, green: no PS loading; red: doubled PS loading; (a) nutrient limitation factor, (b) organic nitrogen, (c) phosphate and (d) dissolved (available) silica.

SWEM models the evolution of DO within the water column as the sum of the divergence of the flux, the rate of change due to production by algae and several oxidation reactions that contribute to the net respiration. These include algal respiration, nitrification, and organic carbon oxidation (bacterial respiration). In addition, in the surface layer there is a gain or loss to or from the atmosphere as a result of atmospheric surface diffusion, and in the bottom layer a loss due to the benthic demand. These rates are not output by SWEM; however, we modified the SWEM model to report these rates by cell and layer at the internal SWEM time step of 96 day^{-1}. Figure 21 shows the evolution of the depth profiles of monthly mean O_2 production (green) and respiration (red) rates at the SWEM grid cell containing station C1. The surface exchange with the atmosphere is shown by the blue line which is nonzero in the surface layer. The model shows weak vertical structure in the respiration rate with maxima at the surface. There is considerable seasonal variation with values reaching 30 μM day^{-1}in the summer.

The respiration rate estimates of Goebel and Kremer (2007) show considerable variability in space and time and so averaging is necessary to reveal coherent patterns. In Figure 22 we show the mean summer (July-September, 1989) respiration rates in the SWEM solution at the top two model cells along a line down the center of the Sound eastwards from the Battery, NY. Also shown are the means and standard deviations of measurements by Goebel and Kremer (2007). It is clear that the field observations are substantially larger, almost double, the values emerging from the SWEM simulations.

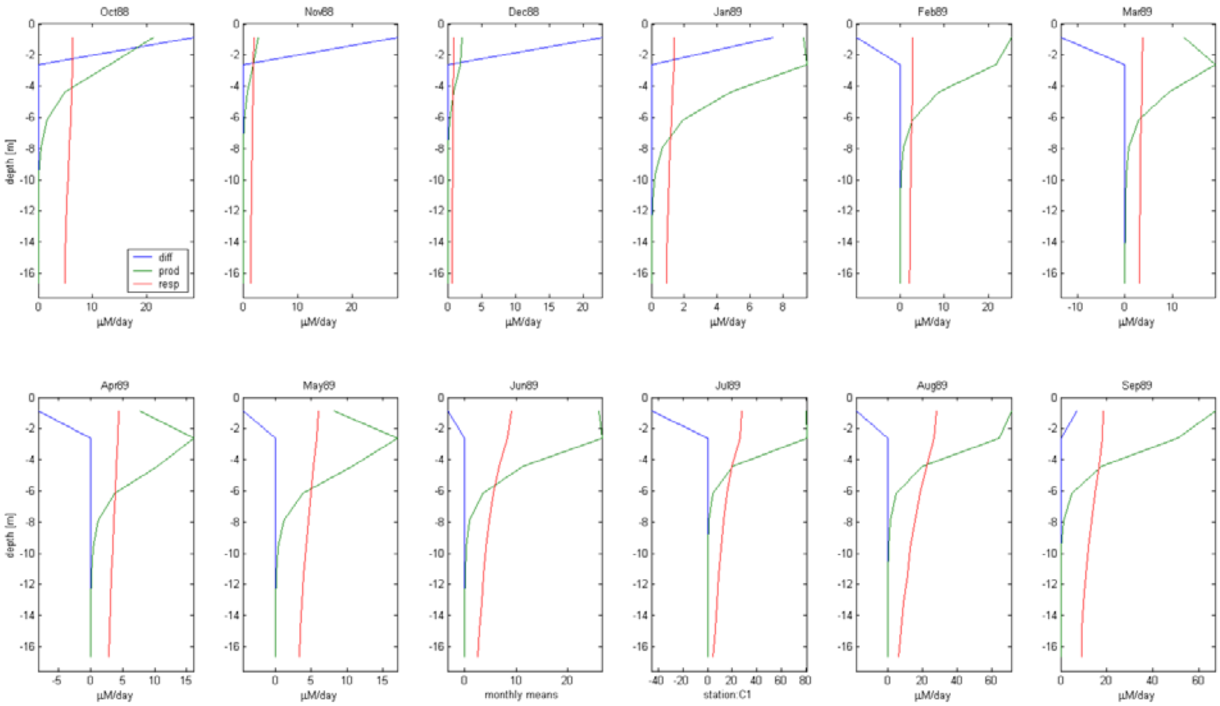


Figure 21. Vertical profiles of the monthly mean O_2 respiration (red) and production (green) rates (µM day^{-1}) by month at station C1 (cell 41,50). The mean net atmospheric exchange at the surface layer is shown by the blue line. Note that it is zero except in the surface layer.

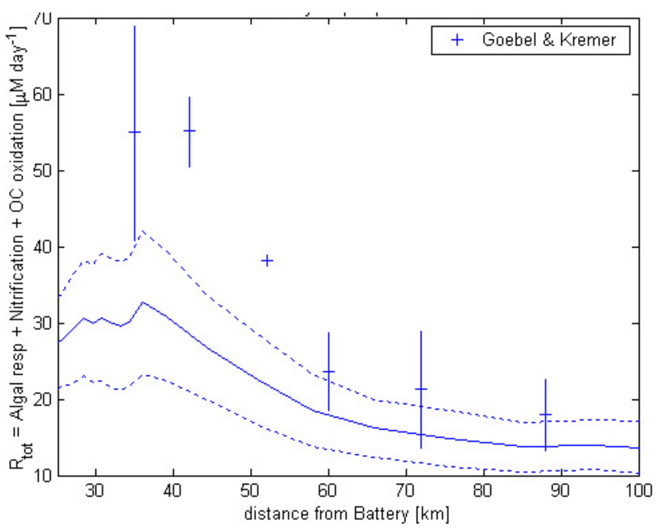


Figure 22. SWEM means (solid horizontal line) and standard deviations (dotted lines) of near surface (layers 1 and 2) July-August O2 respiration rates (µM day─1) plotted by along-sound distance from the NYC Battery. Also shown by the vertical lines are the means and standard deviations of measurements by Goebel and Kremer (2007).

The high-frequency output feature we installed in SWEM also allowed us to observe the short time-scale DO variation in the surface layer of the SWEM model due to the day-night diel photosynthesis cycle. Observations suggest that diel variation in DO in LIS is far greater than what we find in the SWEM solutions. Figure 23a shows the July SWEM model DO cycle in the surface layer in the SWEM grid cell containing DEP station C2 whereas Figure 23b shows observations from 2004 from the nearby LISICOS Execution Rocks buoy. Note that although the SWEM model is consistent with the mean DO levels, it fails to capture the magnitude of the high frequency variability. This suggests that, in addition to the substantial underestimation of the respiration rate commented upon above, SWEM may also be significantly underestimating production rates

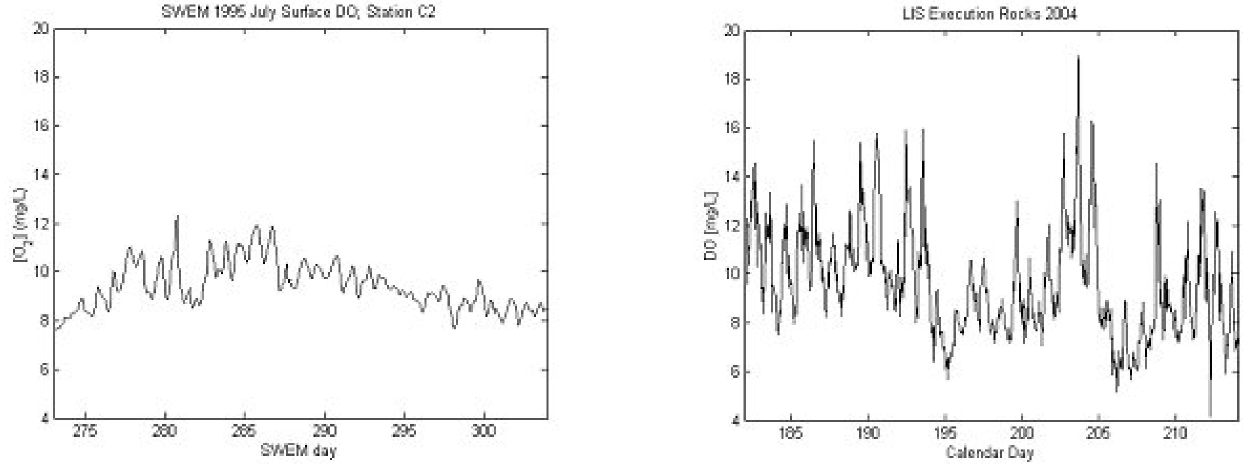


Figure 23. (a) High frequency variation in DO at C2 in the July 95 SWEM simulation and (b), an example of the evolution at a buoy in the same vicinity (Execution Rocks) in July, 2004.

That SWEM substantially underestimates production and respiration rates in LIS provides a candidate explanation for why Hydroqual needed to effectively eliminate the vertical flux of O_2 from the surface to the bottom caused by turbulent mixing. It is possible that low respiration and low mixing could result in the same DO as high respiration and high mixing. Why the respiration in SWEM is too low is an important question that we now address.

Our comparisons of the SWEM simulation with observations of the summertime DO concentrations in Long Island Sound with realistic levels of vertical mixing revealed, as Hydroqual discovered, that the model significantly over-predicts the near bottom DO concentration. To isolate why this occurs, we quantitatively investigated the mechanisms that determine production and respiration rates in SWEM. As has been pointed out above, the respiration rates represented by SWEM are, approximately, a factor of two lower than recent observations indicate they should be. Here we examine what internal parameters affect these rates and suggest how these may be modified.

7.6.3 Sensitivity to the O_2 Budget Formulation

In SWEM the respiration is the sum of algal respiration, nitrification, and organic carbon oxidation (bacterial respiration.) Nitrification represents an insignificant portion of the total. Algal respiration in SWEM is only proportional to the rate of algal mortality. Algal mortality is the product of the algal death rate and the algal population concentration. With this model, therefore, the only way to increase respiration is to increase either the death rate or the population level. However, an increased death rate requires an increased growth rate to sustain the population size. Since either high algal growth or a high algal population requires nutrients (which we discovered by experiment are limiting in SWEM simulations in the summer), neither of these options succeeds in increasing respiration by the amount required. This formulation of the DO budget has the consequence that nutrient limitation effectively caps both the growth rate and the respiration rate. This leaves organic carbon respiration as the only means by which respiration can be increased within SWEM to the levels measured by Goebel and Kremer (2007).

Just as algal respiration in the DO budget is linked to the algal mortality, production of organic carbon and DO in SWEM is proportional to the algal population level and cannot be increased without either increasing the algal population or increasing the growth rate. Either of these options would require unrealistically high nutrient levels. Though we could introduce an alternative formulation of the budget to the SWEM code, testing and documenting that is beyond the scope of the project. However, we can demonstrate the effect of changes in the existing model by modifying the interpretation of the parameter $FLOCEX$, the portion of algal growth going to organic carbon exudation. In SWEM the growth of the algal population biomass is represented as the product of a growth rate and $(1 - FLOXEX)$. If both the algal growth rate and $FLOCEX$ are increased such that the product is unchanged, then production can be increased without effecting the algal population dynamics. Increasing $FLOCEX$ then increases the amount of organic carbon available for DOC respiration, and can thereby increase the total respiration. Of course this is a distortion of the ecosystem model; however, it is a convenient way to demonstrate the sensitivity of the predictions of the model to the details of the DO budget formulation.

During this exercise we also noted a problem with the formulation of the exchange of DO at the surface. In SWEM this is modeled as the product of the atmospheric exchange piston velocity and the difference between the $[O_2]$ concentration and the $[O_2]$ saturation concentration. The atmospheric exchange piston velocity, k_L is calculated as the maximum of a value specified at the cell or a fixed minimum value, $KLMIN$. The SWEM code also had an option to calculate k_L empirically using the near-surface velocity shear, but the shear was calculated incorrectly and the code for this was commented out in more recent versions of the source code.

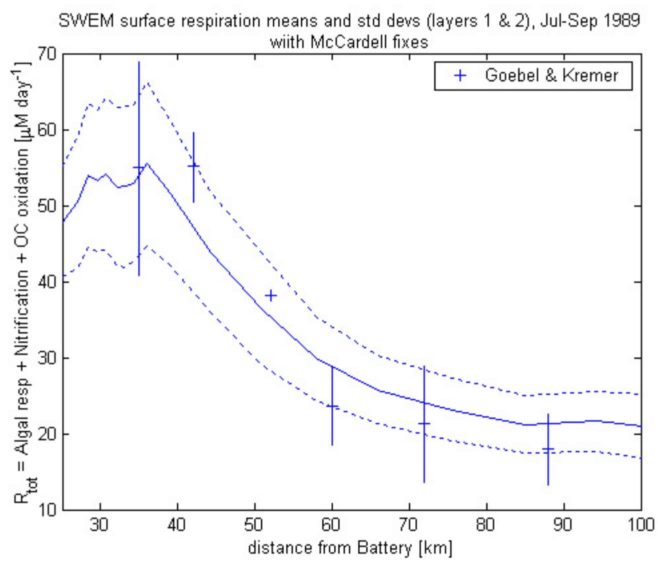


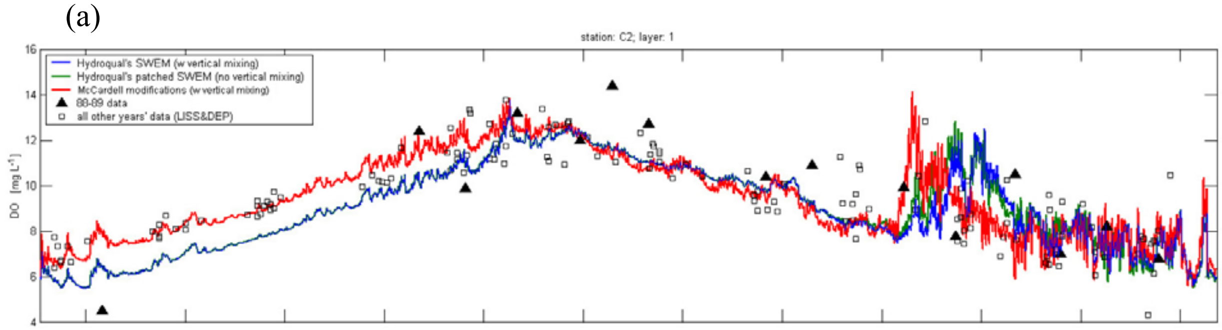
Figure 24. Modified SWEM means (solid horizontal line) and standard deviations (dotted lines) of near surface (layers 1 and 2) July-August O2 respiration rates (μM day─1) plotted by along-sound distance from the NYC Battery. Also shown by the vertical lines are the means and standard deviations of measurements by Goebel and Kremer (2007). Compare to Figure 22.

To evaluate whether the suppression of respiration through its link to the algal biomass production could explain the failure of SWEM to simulate the low DO in the western Sound without unrealistically limiting the rate of vertical mixing in SWEM, we performed experiments which produced more respiration without reducing or increasing the algal biomass. The value of $FLOCEX$ used by Hydroqual is 0.200. We conducted a series of simulations in which we increased this and also increased the algal growth rates such that $growth\ rate(1 - FLOCEX)$ would remain unchanged. This approach increased the high frequency DO variance and made the near surface DO concentration more sensitive to the value of k_L. The value used by Hydroqual for the minimum piston velocity, KLMIN was 0.300 m day^{-1}. Recent studies (Marino and Howarth, 1993; Raymond and Cole, 2001; Zappa et al, 2003) indicate that this value is quite low and that $1-3$ m day^{-1} is a more appropriate value. We used values in this range.

Figure 24 shows a comparison of the respiration rates using the modified SWEM model with vertical mixing reinstated and higher production and respiration. The rates measured by Goebel and Kremer (2007) are shown by the vertical bars for comparison with the SWEM values. Recall that the original SWEM results are shown in Figure 22. Clearly it is possible to obtain substantially higher respiration rates in the modified version of SWEM.

The corresponding evolution of the DO in the surface and bottom cells at C2 in the 88-89 simulation is shown in Figure 25. The results from the original model are shown by the green lines, the original model with vertical mixing reinstated by the blue lines, and the version of the model with realistic mixing and enhanced production and respiration is represented by the red lines. We employed $FLOCX = 0.5$ and increased the maximum phytoplankton growth rate parameters to $K1C = 2.72$ and $K2C = 4.8$ to ensure that the algal and chlorophyll concentration remained the same. We also increased the surface exchange by choosing $KLMIN = 2.5$. In the surface layer (Figure 25a) the modified SWEM (red line) is in much better agreement with the observations throughout the fall and winter than the original SWEM (blue line). In mid-June differences in the models become apparent in the lower layer. The SWEM version with mixing is anomalously higher than the observations, the fact that led Hydroqual to distort the hydrodynamics. The green and red lines appear to be consistent with the survey data though the modified SWEM version exhibits the lowest DO values.

(a)



(b)

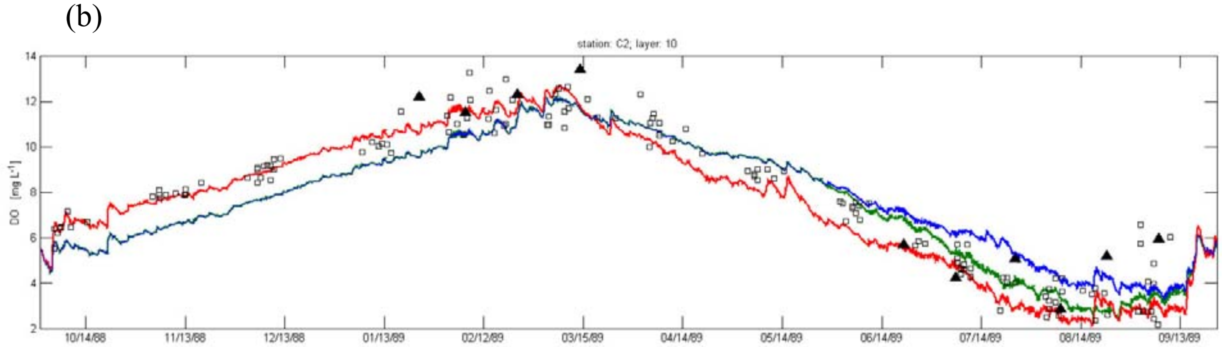


Figure 25. : Comparison of the time evolution of SWEM DO concentrations (mg l^{-1}) at the surface (a) and bottom (b) at C2 in 1988-89. The results obtained with the original SWEM (with vertical mixing reduced) are shown in green. The model with vertical mixing reinstated but otherwise unmodified is represented by the blue lines and the model with our modifications is shown by the red line. The solid triangles represent observations made by the CTDEP at station C2 for the model run year. The hollow squares represent all observations made by the CTDEP at station C2 for all other years through 2005.

Though the differences in the actual concentrations are quite small, the effect of the high bias in SWEM on the statistics used to evaluate the impact of management actions is substantial. In Figure 26 the blue line shows the number of days that the concentration of DO at station C2 fell below the value on the abscissa in the 1988-89 simulation. Figure 26 shows that the DO fell below 3.5 mg/l for only a few days using the original SWEM model but the modified model results show hypoxia for almost 10 days.

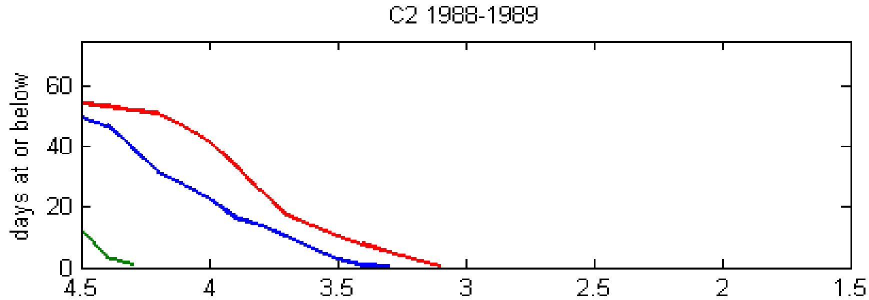


Figure 26. Number of days that the solution at station C2 falls below a threshold. The blue line represents the solutions obtained with the original SWEM code. The green curve shows the SWEM solution with mixing reinstated, and the red line show our modified model that includes increased production, respiration and vertical mixing.

7.6.4 Summary of Sensitivity to Model Formulation

It is evident that SWEM is very sensitive to the formulation of the model. This is especially true for simulations of hypoxia metrics like duration and area below a threshold value. We have demonstrated significant changes when using alternate formulations. The absence of mixing in SWEM is clearly a distortion of the physical processes in LIS that was motivated by a desire to better represent the seasonal evolution of hypoxia. However, we are confident that this ad hoc adjustment to reduce vertical mixing fluxes masks a serious deficiency in the biogeochemistry model. We have demonstrated an approach that provides preliminary guidance on how this problem can be fixed. Moreover, these improvements would allow the model to represent variability in physical transport processes that were eliminated from the model by Hydroqual.

7.7 Tasks 7 and 8. Simulations for the years 1999-2002

The validation and verification studies of SWEM were conducted by Hydroqual Inc. using data for the years 1988-89 and 1994-95. However, it is well established that there are significant differences in the pattern of river discharge and wind variations from year to year, so it is important to assess whether those simulations are representative. As part of this project, we subcontracted with Hydroqual to compute the circulation or hydrodynamic fields for four additional years (1998-1999, 1999-2000, 2000-2001, and 2001-2002) and used the resulting transport coefficients, together with the 1994-95 boundary source fluxes, to simulate what the conditions could have been in the Sound had weather conditions been different. We also repeated these simulations with the original model with mixing restored and with the modified model in which respiration and production were enhanced to values consistent with observations and with mixing restored.

Using the original SWEM code and the 1994-94 source data, we performed year-long simulations with the 6 sets of circulation fields. The solutions for the near bottom DO at DEP station C2 are shown by the blue lines in the six panels of Figure 27. The black triangles show the measurements for the year of the circulation forcing and the open squares show the rest of the near bottom DO measurements near C2 in the archive. Obviously, evaluation of the skill and comparison of the values obtained for 1988-89 and 1994-95 for these simulations is not informative since the discharges were likely different in each of the years. It is interesting that the variation between the simulations is comparable to the variation in data. This is more clearly demonstrated in Figure 28 which shows the evolution of the standard deviation in the observations and the simulations. This suggests that the influence of inter-annual fluctuations in forcing of the circulation and mixing is significant.

The green lines in Figure 27 show the SWEM solutions when the ad-hoc reduction in the turbulent mixing rates is removed. These verify that the SWEM predictions would be well above the observed DO concentrations for a wide variety of circulation patterns. The red lines show the solutions obtained with the modified model in which respiration and production were enhanced and vertical mixing was returned to realistic levels. These show the lowest near bottom DO in all years. Note that we did not attempt to "calibrate" this model since the biogeochemistry is distorted and our intent was simply to illustrate the significance of the choices implicit in the selection of the model formulation.

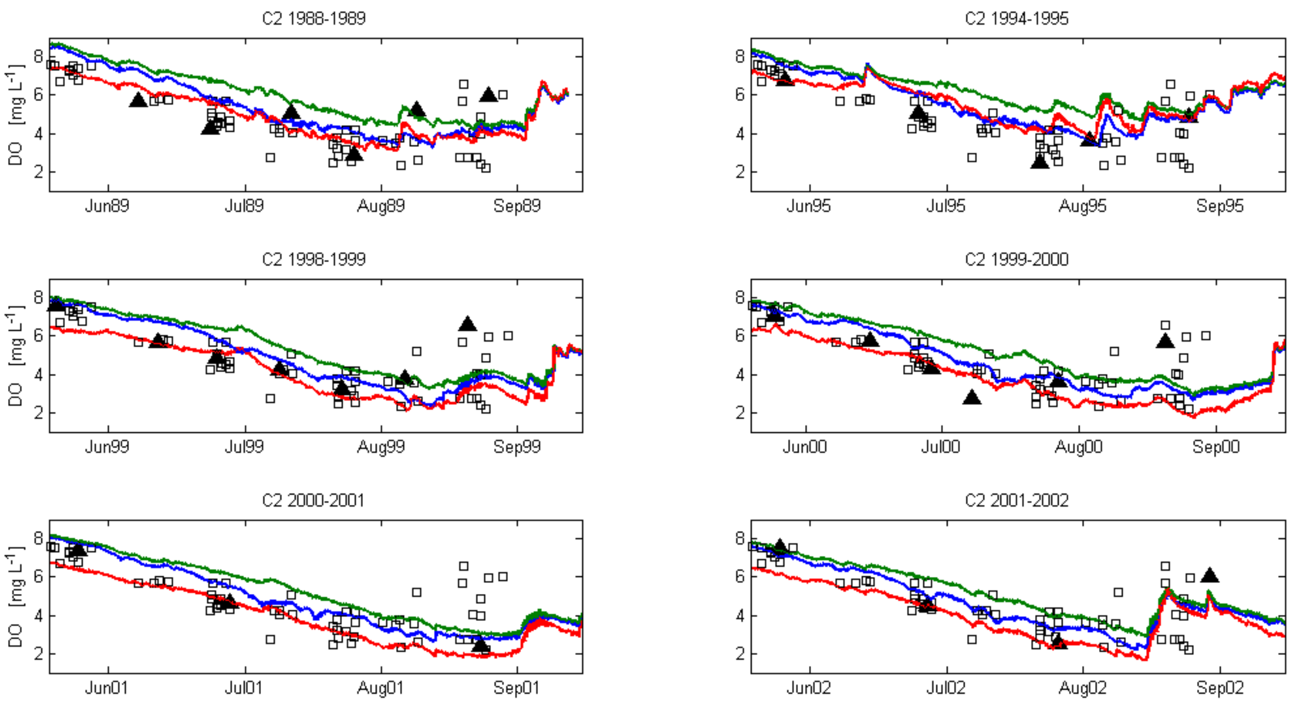


Figure 27. The solutions for the near bottom DO at DEP station C2 using the original SWEM code (blue lines). The black triangles show the DO observations for the year of the simulated circulation and the open squares show the rest of the near bottom DO data in the archive for C2. The solution using the original model with mixing restored is shown by the green line and the solution using the modified model with enhanced production and respiration, and realistic mixing is shown by the red line.

Figure 28 shows the mean evolution of the unmodified simulations and the data and demonstrates that the near bottom DO simulations at B3 and C2 are biased high by approximately 1 mg/l between June and August. Analogous graphs demonstrate that this is true at other western Sound sites as well. It is also worthy of note that the 1988-89 and 1994-95 years studied by Hydroqual had forcing fields that led to the highest summer time DO concentrations of the 6 years examined.

The effect of this bias on the statistics used to evaluate the impact on the duration of hypoxia statistic is illustrated in Figure 29 in which the blue lines show the number of days that the concentration of DO at station C2 fell below the value on the abscissa as in Figure 26. In the predictions for summer of 1989 and 1995, the DO fell below 3.5 mg/l for only a few days. In the other 4 years, the DO fell below this threshold for more than 20 days. The green lines show the predictions for the original SWEM model with mixing included and the red lines show the results for the modified model. Clearly the modified model predicts much longer periods of hypoxia. The recent analysis of the DO archive (The 2009 Season review, O'Brien-

Clayton) by the CTDEP presented a summary of the observed variation in the duration of hypoxia and reported the average for the 3.0 mg/l standard as 56 ± 14 days. In none of the simulations with the original SWEM code does the duration of hypoxia exceed 25 days. It is evident that DO in the model is biased high and that this has a major influence on its ability to predict the duration of hypoxia.

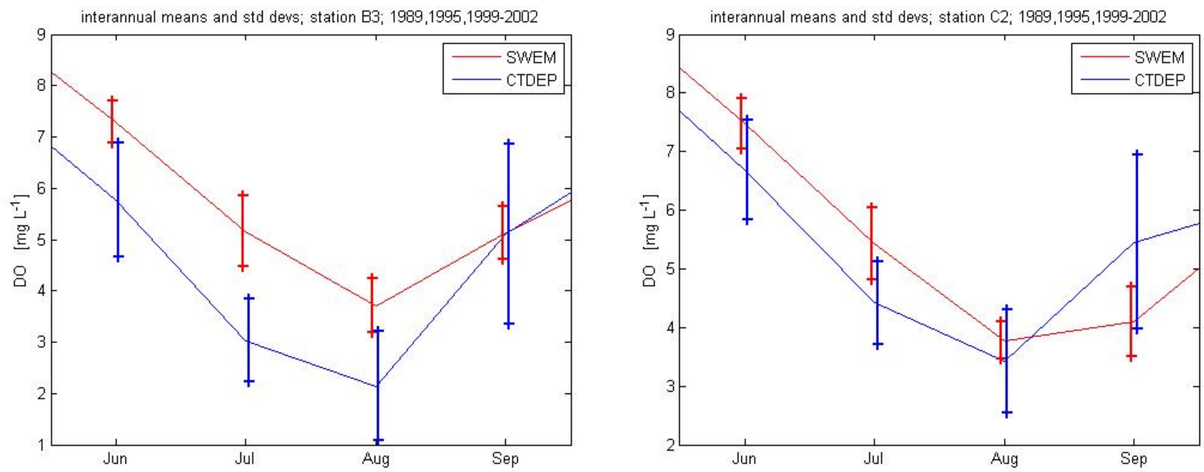


Figure 28. Time series of the means and the variation between years of the observations at B3 (left panel) and C2 (right panel) in the 6 years shown in Figure 27 (blue) and the corresponding means and variability of the (unmodified) model simulations of these 6 years (red).

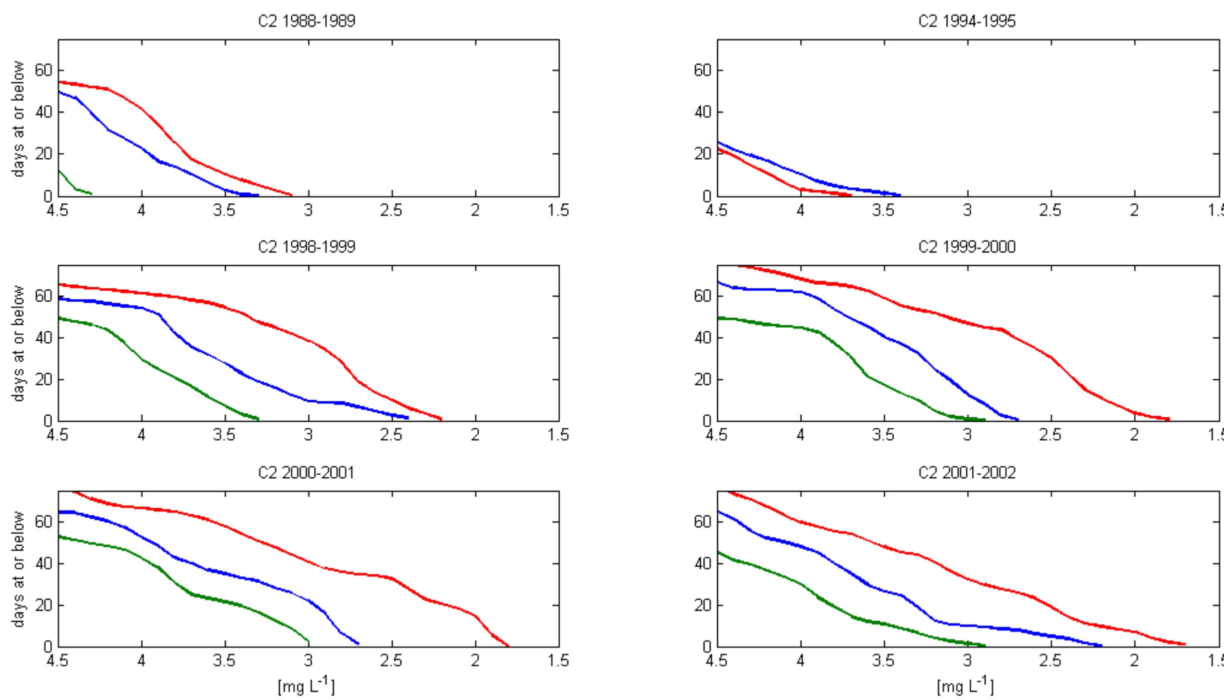


Figure 29. Duration the solution at station C2 falls below a threshold. The blue line represents the solutions obtained with the original SWEM code. The green curve shows the SWEM solution with mixing reinstated, and the red line show our modified model that includes increased respiration and vertical mixing.

Task 9. Short-term (monthly) forecasts of the extent of the hypoxia in the summer.

In the original proposal we anticipated the SWEM forecasts would have demonstrable skill in forecasting the extent and duration of hypoxia. However, the analysis of the skill parameter S_1, defined in Equation 1, demonstrated that the SWEM forecasts were less informative than the climatology of monthly means. Figure 30 shows a summary of the skill for DO in the 6 years we examined for the original SWEM model and our modified model. Though the modifications improved the skill significantly, the values are all still negative. The best predictor of the future is therefore the past climatology.

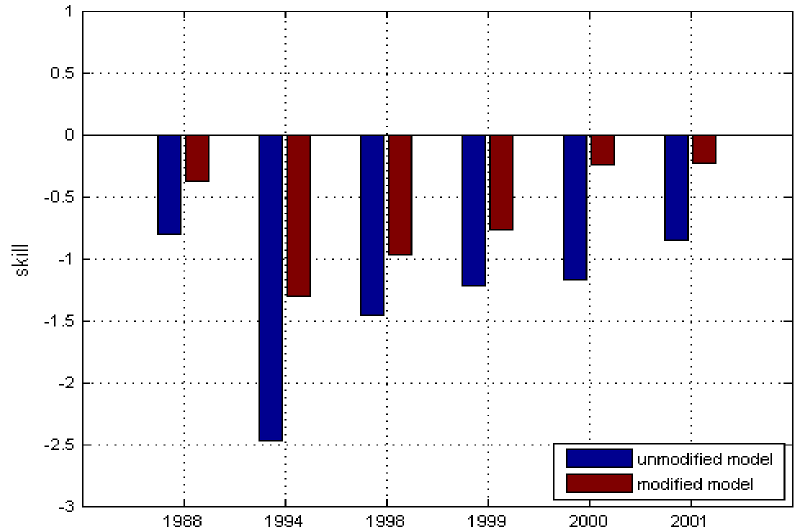


Figure 30. Skill for the DO variable for the 6 years simulated using both the unmodified version of SWEM and the version with enhanced production and respiration, and realistic mixing.

Task 10. Development of a Tool to Guide Model Improvement.

We believe that the use of the Brier skill, S_1, as defined in Equation (1), is a valuable metric for determining which model modifications lead to quantitatively better simulations. We have used the climatology established from the CTDEP surveys as the reference model in this project but the Hydroqual version of the SWEM model could be used instead so that S_1 would indicate whether new simulations were improvements or not. Further, the skill of components of the model can be assessed by examining the skill for individual variables. Though the Brier skill of DO levels is most significant for the purposes of TMDL development, the fact that the skill of the simulations of some variables is large and negative (e.g. the dissolved silica and phosphate in the 1994-5 simulations as shown in Figure 6) suggests that the representation of the cycles of these nutrients requires more attention. The negative skill could results from bias in the observations, inadequate calibration, or missing processes in the model. Whatever the cause, quantitative measures of skill provide guidance to both scientists and managers on how to invest in further research and on whether adding new components to the model is likely to lead to real model improvement or just a more complicated model. During the development of SWEM it was judged to be better to prescribe the level and seasonal variation in the zooplankton grazing rate rather than model the dynamics of the population variation. This type of decision could be addressed more quantitatively using S_1. We could, for example, prescribe the dissolved silica and phosphate variation and assess the effect in the DO skill. If the skill is higher, then this simplification might be worthwhile.

Task 11. Recommendations

The keys to a quantitative understanding of the links between the management of nitrogen discharges and the duration and extent of hypoxia are:

(1) The development of a testable and tested model of the processes that link nutrient discharges and the respiration of organic matter in the western Sound, and

(2) The development of a testable and tested model of the vertical transport of oxygen in the western Sound.

We believe we have demonstrated that the version of SWEM that we examined requires an unrealistic ad-hoc reduction in the turbulent transport rates of dissolved oxygen, and all other dissolved materials, in order to establish hypoxic conditions in the near bottom waters of the western Sound in the summer and we have shown this may be a consequence of an inadequate formulation of the production and respiration components of the model.

We recommend that we must immediately establish the sensitivity of the management decisions that have been made using SWEM to the misrepresentation of vertical mixing. A few minor changes to the existing SWEM and a calibration effort with existing data of the type we outlined in Section 7.2 would allow this to be accomplished quite quickly.

To move ahead to establish a more effective prediction capability we believe that mixing must be restored to the model and that substantial modifications to the production and respiration rate representations must be introduced and tested directly. By that we mean the rates must be measured and compared to the model predictions. The same is true for the gas transfer rates across the air -sea surface and across the pycnocline. Comparing predictions to ship surveys of DO concentrations is unlikely to distinguish between alternative formulations or parameter choices. The work of Goebel and Kremer (2007) has established the magnitude of production and respiration and their variability. A careful and sustained measurement campaign that resolves the spatial and temporal variability of these processes is a key step to the development of an improved model. To understand the variability in the respiration and production and the link to nitrogen, a more extensive field campaign may be necessary but the wise development of a plan for that must await an appropriate depiction of the variability.

Though the circulation and transport in SWEM was not a substantial limitation, modern models with much higher resolution of the coastal geometry and bathymetry are now available that would allow better representation of the effect of the hypoxia in small bays and inlets. Since use could be made of an existing community model, this should not be an onerous task, and a revised water quality model, or models, should consider adopting a new transport component.

Our analysis of the skill of SWEM in predicting the various variables showed that some elements were much better simulated than others. The simulation of nitrate and ammonium were particularly ineffective. Whether this is due to poor model performance or to large uncertainty in the data due to a high degree of natural variability is unclear. More measurements at high frequency must be obtained to properly resolve this issue. We note that zooplankton populations were not included in SWEM. A rational approach to deciding what should be simulated and what should be prescribed based on quantitative skill assessment should be developed.

The effort that we have expended in learning how to interact with SWEM has provided valuable lessons that are important to record. The unusual data formats used in SWEM are very inconvenient and considerable computing effort is required to extract variables for analysis and comparison to observations. We recommend that future modeling efforts require the use of community standards for saving solutions. Standards such as NetCDF Climate and Forecast (CF) metadata conventions for the archiving and sharing of model solutions should be adopted. The limited documentation and the proprietary code is also an obstacle for model evaluation and improvement. We recommend future efforts reject proprietary code and insist on open-source model development with modern source and revision control standards. There are several such models with documentation and analysis tools available. These steps would allow more interaction between models and observations which would undoubtedly accelerate the development of improved models.

References:

Bennett, D. M., J. O'Donnell, W.F. Bohlen and A.E. Houk (2010). Tides and Overtides in Long Island Sound. Journal of Marine Research, J. Marine Res. (in press)

Dillon, T.M.,1982: Vertical overturns: A comparison of Thorpe and Ozmidov length scales. J. Geophys. Res., 87, 9601-9613.

Goebel, N.L. and Kremer, J.N., 2007. Temporal and Spatial Variability of Photosynthetic Parameters and Community Respiration in Long Island Sound. *Marine Ecology Progress Series*. Vol. 329: 23-42.

Hydroqual Inc. (2003) Presentation to the Long Island Sound Science and Technical Advisory Committee, Stony Brook University, October 23-24, 2003.

Marino, R., Howarth, R.W., 1993. Atmospheric Oxygen Exchange in the Hudson River: Dome Measurements and Comparison With Other Natural Waters. *Estuaries*, Vol. 16, No. 3a.

McCardell, G. and O'Donnell, J., 2009: A novel method for estimating vertical eddy diffusivities using diurnal signals with application to western Long Island Sound. *J. Mar. Syst*., doi:10.1016/j.jmarsys.2008.10.011.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20**, 851–875.

O'Donnell, J., Dam, H.G., Bohlen, W.F., Fitzgerald, W., Gay, P., Houk, A.E., Cohen, D.C., Howard-Strobel, M.M., 2008. Intermittent Ventilation in the Hypoxic Zone of Western Long Island Sound During the Summer of 2004. *Journal of Geophysical Research* 113, C09025. doi:10.1029/2007JC004716

Odum, H. T., 1956: Primary production in flowing waters. *Limnology and Oceanography* 1(2).

Osborn, T. R., (1980) Estimates of the local rate of vertical diffusion from dissipation measurements. J. Phys. Oceanogr., 10, 83–89

Peters, H., 1996: Observations of stratified turbulent mixing in an estuary. Neap-to- spring variations during high river flow. Estuarine, Coastal and Shelf Sci., 45(1), 69-88

Ruddick, B, A. Anis, and K. Thompson (2000) Maximum Likelihood Spectral Fitting: The Batchelor Spectrum. J. Atmos. And Oceanic Tech. 17, 1541-1545.

Thorpe, S. A., (1977) Turbulence and mixing in a Scottish loch. Philos. Trans. Roy. Soc. London A, 286, 125–181.

Thorpe, S.A. (2005) *The Turbulent Ocean*. Cambridge University Press.

Raymond, P.A., Cole, J.J., 2001. Gas Exchange in Rivers and Estuaries: Choosing a Gas Transfer Velocity. *Estuaries*, Vol. 24, No. 2.

Zappa, C.J., Raymond, P.A., Terray, E.A., McGilis, W.R., 2003. Variation in Surface Turbulence and the Gas Transfer Velocity over a Tidal Cycle in a Macro-tidal Estuary. *Estuaries*, Vol. 26, No. 6.

8. Summary of Findings

Low levels of dissolved oxygen (DO) occur in the near bottom waters of western Long Island Sound (LIS) each summer which stresses the marine life in the region. This phenomenon occurs in many urbanized estuaries and is termed hypoxia. The EPA, and the States of New York and Connecticut, have developed a Comprehensive Conservation and Management Plan (CCMP) to mitigate the extent and duration of hypoxia, using a computer model to assess the likely impact of reductions in nitrogen discharged from water treatment plants and non-point sources.

An improved model, the System Wide Eutrophication Model (SWEM), has been developed to simulate the biogeochemistry and circulation in Long Island Sound and adjacent waters and is being used to reassess the effectiveness of the plan. SWEM is a complex model with many parameters that represent the rates of the processes that influence the DO concentration. In this project we implemented SWEM at the University of Connecticut to independently study the sensitivity of the model predictions of DO concentrations to parameter choices and boundary conditions, and to assess the effect of year-to-year variations in precipitation and wind patterns. Since there have been many important advances in the understanding of LIS and the coastal ocean since the initiation of the development of SWEM, we also assess whether SWEM is consistent with latest knowledge.

With the support of the EPA and funds from the University of Connecticut, we have constructed a 48-processor parallel computing system with 2 Terabytes of disk and installed the SWEM code. This capability enables rapid calculations and the simulation of a wide range of model scenarios with which to assess the sensitivity of the predictions to parameter choices. We also built a MATLAB-based system to access the solution files, perform analyses and generate graphics. To ensure that our implementation of SWEM was performing properly, we replicated the simulations of Hydroqual Inc. for 1995 and demonstrated good agreement.

To quantitatively assess how well the model performed we employed a statistical summary of the misfit between the model solution and water quality observations made by ship surveys in 1988 and 1995. We called this the *skill*. We used several variants of this metric, but for the purposes of comparing different versions of the same model, a higher skill means a better simulation. SWEM has 120 parameter values that must be selected. We examined the effect on the model skill of increasing and decreasing each parameter by 10% from the values chosen by Hydroqual. This analysis required 240 simulations for each of 1998 and 1995. The results demonstrated which parameters were most important and it suggested some changes to parameter values that might lead to improved predictions.

We also investigated the consequences of doubling the fluxes of nitrogen from point sources and setting these to zero on the predictions of the model. We found little difference between the solutions in areas prone to hypoxia. With zero discharges, the minimum DO in the late summer was approximately 1mg/L higher than with the 1995 discharge levels and only 0.1 mg/L lower with double the 1995 discharges. Since SWEM was designed to mimic the existing plankton community structure, which would likely be different if these ratios were radically different, these simulations should not be considered realistic. However, with the caveat that the mixing, respiration and production rates are all probably too small, they provide a bound on the expected magnitude of DO changes to be expected from management actions which might reduce discharges 30 to 50%. The changes we expect are likely to be small.

The developers of SWEM found it difficult to establish hypoxic conditions in the model solutions and resorted to an ad-hoc adjustment to the vertical mixing rates predicted by the circulation/hydrodynamic module. Recent work on mixing in the coastal ocean suggests that the reductions they imposed were unrealistic and that the original values were reasonable. This suggested to us that the need for the distortion of mixing masked other under-lying problems in SWEM.

A recent field program, also supported by the EPA, has revealed the magnitude of the community production and respiration in LIS that SWEM should predict. Another study maintained buoys in the LIS to measure the high frequency variability of the DO near the surface, in the pycnocline and near the bottom, and made some preliminary estimates of vertical mixing rates in the western Sound. Since measurements of this type were not available during the model development, SWEM did not explicitly compute and report these quantities. We modified the SWEM code appropriately and compared these results to observations. We found that both respiration and production were significantly underestimated in SWEM and that the levels of mixing predicted by the original circulation module were in the realistic range. We conclude that it is likely that the SWEM simulations required vertical mixing to be reduced in order to match the observed trends in summertime near-bottom DO in the western Sound because respiration was too small.

Although outside the scope of the original project, we investigated the cause of the under-estimation of respiration. We performed some numerical experiments to check that increasing the respiration was a quantitatively important change and demonstrated that it made a significant difference. Without the mixing distortion and with a few other parameter changes, increasing the respiration improved the model skill.

The validation and verification studies of SWEM were conducted by Hydroqual Inc. using data for the years 1988-89 and 1994-95. However, it is well established that there are significant differences in the pattern of river discharge and wind variations from year to year, so it is important to assess whether the simulations are representative. We contracted with Hydroqual to compute the circulation (or hydrodynamic) fields for four additional years (1998-99, 1999-2000, 2000-2001, and 2001-2002) and used these, together with the 1994-95 boundary source fluxes, to simulate what the conditions in the Sound would have been with different weather conditions. We then compared these solutions and found that they were similar to each other. We conclude that we don't have the ability within SWEM to distinguish the influence of inter-annual meteorological variability on hypoxia.

The circulation/hydrodynamic module in SWEM is used to compute the horizontal transport and vertical mixing rates of the water quality variables. Since new data was available to us, we used it to evaluate vertical mixing rates and the horizontal velocity predictions. We found that this component of the model did quite well and the original vertical mixing rates were not inconsistent with the observations. The main limitation is that the lateral (almost north-south) variation of the circulation is not resolved because the resolution of the model is crude. In most of the hypoxic areas there are only 1 to 5 grid points across the Sound. Since the solution technique smoothes the fields, there is little lateral structure in the simulations. Recent work suggests that this may be an important weakness in the current model formulation.

We conclude our study with recommendations for future research needs. Clearly, the model must be improved to better represent the critical processes that determine the rate of decline in DO in the summer: respiration and vertical mixing. We have outlined how that should be accomplished. This is a substantial change and the implications on the management decisions that have been instituted should be evaluated by repeating at least some of the simulations used in the development of total maximum daily loads (TMDLs). However, that is not enough. For the model to be useful in predicting the impact of changes to sources of nutrients, it is also essential to observe how community respiration and production rates depend upon the available nutrient concentrations in LIS. The simulation of that response will establish confidence that the model can represent the effect of management actions correctly. This will require sustained field observations in addition to the existing monitoring program.

Recent observations have shown that the magnitude of the variation in DO and chlorophyll that occurs during a single day is comparable to the differences in the monthly mean concentration between years. It seems likely that nutrients will show similar variability. Since the ship surveys occur during the day and only twice a month, it is not surprising that comparing this data to the model predictions does not really assess whether the model is adequate. Some of the differences between the predictions and observations are due to the uncertainty in the observation of the several-day mean. Though the precision of the measurements may be high, the sampling scheme does not provide the best estimate of the values that SWEM was being used to predict.

Further, only the availability of high frequency moored DO observations and ship based measurements of respiration and production rates provided the insight that led to identification of the weaknesses in the model. The existing observations of the community respiration and production rates are highly variable in time and location. It is unclear what causes this variability. The high frequency variability in the nutrient levels are unknown at present and this must be established. Finally, the crude resolution of the lateral variation in the circulation model essentially prohibits the possibility that lateral gradients strong enough to drive substantial lateral circulation can be established. The low spatial resolution also diminishes the model's value to the management of water quality in embayments like Hempstead Harbor, Smithtown Bay, etc.

We believe that some improvements to the modeling process are needed. It is critical that engagement between observationalists and theoreticians be enhanced. Maintaining an open source model code that is freely accessible using community standards for the integration and documentation of software updates is essential to allow broad participation in the assessment and development of the model. It is essential that community data standards such as NETCDF Climate and Forecast (CF) metadata conventions for the archiving and sharing of model solutions should be adopted. These conventions are designed to promote the processing and sharing of files. This will enable more independent and critical evaluations to take place. Further, the maintenance of an open access data system with observations of both forcing fields and data for evaluating solutions in a standard format will facilitate the evaluation of models and model variants. Thredds (Thematic Real-time Environmental Distributed Data Services) data servers provide an interoperable platform for accessing model output and allow for streamlined data integration into analysis tools. This infrastructure will both accelerate the integration of new knowledge and enhance community support for the results.

The forecasting of uncertainty in predictions should be required in the future. This study initiated that process but because we discovered several fundamental weaknesses in the model, it must be repeated with an improved model. The fields of weather forecasting and climate change prediction have adopted the approach that several models with alternative parameterizations of model processes should be used in policy decision making. This allows the uncertainty in science to be quantitatively projected into the policy formulation. Our proposal that an open source modeling system be developed is consistent with this idea since it would allow modules to be modified by scientists, and the impact of different models assessed. Finally, we believe that models should be as simple as possible, and additional processes should only be incorporated when there is quantitative evidence that they improve the skill.

Acknowledgements

We are grateful to numerous colleagues for sharing insights that led us to better understand the SWEM model and the biogeochemistry of Long Island Sound. Bill Fitzgerald, George McManus, Jim Kremer, Paul Stacey and Senjie Lin were particularly helpful. We are also grateful to Jim Fitzpatrick and Robin Landeck Miller of Hydroqual Inc. who provided essential advice and technical support in the early stage of our project. The availability of observations from the LISICOS buoys (supported by the NOAA IOOS program) and the ship surveys conducted by the Connecticut Department of Environmental Protection, provoked the important insights reported here. These would not have been available without the skill and efforts of the UConn LISICOS team (Frank Bohlen, Adam Houk, Kay Howard Strobel and John Bayruther) and Matt Lyman and Christine Olsen of the CT DEP. This project was supported by funds from the EPA Long Island Sound Program and the University of Connecticut.